



OPEN

Differentiation between multiple sclerosis and neuromyelitis optica spectrum disorder using a deep learning model

Jin Myoung Seok^{1,9}, Wanze Cho^{2,9}, Yeon Hak Chung^{3,4}, Hyunjin Ju^{3,4}, Sung Tae Kim⁵, Joon-Kyung Seong^{2,6,7}✉ & Ju-Hong Min^{3,4,8}✉

Multiple sclerosis (MS) and neuromyelitis optica spectrum disorder (NMOSD) are autoimmune inflammatory disorders of the central nervous system (CNS) with similar characteristics. The differential diagnosis between MS and NMOSD is critical for initiating early effective therapy. In this study, we developed a deep learning model to differentiate between multiple sclerosis (MS) and neuromyelitis optica spectrum disorder (NMOSD) using brain magnetic resonance imaging (MRI) data. The model was based on a modified ResNet18 convolution neural network trained with 5-channel images created by selecting five 2D slices of 3D FLAIR images. The accuracy of the model was 76.1%, with a sensitivity of 77.3% and a specificity of 74.8%. Positive and negative predictive values were 76.9% and 78.6%, respectively, with an area under the curve of 0.85. Application of Grad-CAM to the model revealed that white matter lesions were the major classifier. This compact model may aid in the differential diagnosis of MS and NMOSD in clinical practice.

Multiple sclerosis (MS) and neuromyelitis optica spectrum disorder (NMOSD) are autoimmune inflammatory disorders of the central nervous system (CNS) that have similar clinical features^{1,2}. A disease-specific autoantibody targeting aquaporin-4 (AQP4 antibody) has been discovered in NMOSD, which can help differentiate NMOSD from MS³. However, the antibody assay has variable sensitivity⁴ and can produce false-negative results^{5,6}; in addition, levels of antibody could decrease during NMOSD remission^{7,8}. In terms of brain magnetic resonance imaging (MRI) lesions in MS and NMOSD, the presence of a lesion in the inferior temporal lobe or adjacent to the lateral ventricle, a U-fiber lesion, or a Dawson's finger-type lesion are more suggestive of MS than NMOSD. In contrast, longitudinally extensive transverse myelitis, extensive hemispheric lesions, and periependymal lesions are observed mainly in NMOSD^{9,10}. Nevertheless, differentiation between MS and NMOSD can still be challenging in specific clinical situations¹¹.

Recent machine-learning algorithms have been applied clinically in various neurological diseases¹². In CNS demyelinating disorders, various aspects of the diseases have been evaluated with machine-learning methods. Some authors have reported promising results applying machine learning methods to user-defined features including clinical characteristics, T2 lesion volume, regional gray matter volume, and regional fractional anisotropy values to differentiate NMOSD from MS¹³. However, only a few studies have applied deep learning algorithms to differentiate NMOSD from MS¹⁴.

In this study, we aimed to develop a compact and robust deep learning model to differentiate MS and NMOSD using brain MRI data and offer visual explanations for the resulting classification.

¹Department of Neurology, Soonchunhyang University Hospital Cheonan, Soonchunhyang University College of Medicine, Cheonan, South Korea. ²Department of Artificial Intelligence, Korea University, Seoul, South Korea. ³Department of Neurology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, South Korea. ⁴Department of Neurology, Neuroscience Center, Samsung Medical Center, Seoul, South Korea. ⁵Department of Radiology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, South Korea. ⁶School of Biomedical Engineering, Korea University, Seoul, South Korea. ⁷Interdisciplinary Program in Precision Public Health, Korea University, Seoul, South Korea. ⁸Department of Health Sciences and Technology, Samsung Advanced Institute for Health Sciences & Technology (SAIHST), Sungkyunkwan University, 50 Irwon-dong, Gangnam-gu, Seoul 135-710, South Korea. ⁹These authors contributed equally: Jin Myoung Seok and Wanze Cho. ✉email: jkseong@korea.ac.kr; juhongm@skku.edu

Results

Demographic and clinical features. Eighty-six patients with MS and 70 patients with NMOSD were finally enrolled in this study; 199 MRI scans (86 baseline and 113 follow-up scans) from patients with MS and 109 MRI scans (70 baseline and 39 follow-up scans) from patients with NMOSD were used for classification modeling (Table 1). MS patients were younger than patients with NMOSD (MS, 35.0 ± 9.9 years; NMOSD, 43.9 ± 12.6 years; $P < 0.001$); at the time MRI scan, most of the MS patients (92.5%) were relapsing–remitting type MS (RRMS). Proportions of females were not significantly different between the two groups (MS, 72.1%; NMOSD, 85.7%; $P = 0.063$). All patients were seronegative for the myelin oligodendrocyte glycoprotein autoantibody (MOG antibody), and most patients with NMOSD (66 of 70, 94.3%) were seropositive for the AQP4 antibody. The neurologic disability at the time of the MRI scans in patients with NMOSD or MS were different; the NMOSD group demonstrated a higher EDSS score compared to that of the MS group (median EDSS score, 2.5 vs. 1.0, respectively; $P < 0.001$).

Conventional MRI findings. Mean disease duration at the time of MRI was 5.4 ± 5.4 years, which differed between the two diseases (MS, 5.8 ± 5.3 years; NMOSD, 4.8 ± 5.4 years; $P = 0.020$). Most MRI scans (74.7%) were performed during remission; 77.9% of MRI scans in MS patients and 68.8% in NMOSD patients were taken during remission ($P = 0.106$). In MRI scans of NMOSD patients, 19.3% ($N = 21/109$) had normal findings, and 80.7% ($N = 88/109$) had abnormal findings. Based on the previous classification, 45.9% ($N = 50/109$) of MRI scans showed NMOSD-specific brain lesions, such as longitudinal corticospinal tracts lesions (10.1%, $N = 11/109$), extensive hemispheric lesions (14.7%, $N = 16/109$), periependymal lesions (38.5%, $N = 42/109$), and cervicomedullary lesions (5.5%, $N = 6/109$)^{15,16}. NMOSD-specific brain lesions were observed in only 38.7% ($N = 29/75$) of MRI scans when we included only those MRI scans obtained during remission. Of the 199 MRI scans from the patients with MS, 164 (164/199 scans, 82.4%) scans met the Barkhof criteria, and 38 (38/198 scans, 19.2%; one MRI was excluded as it was taken without enhancement) scans showed T1 enhancing lesions; 25.0% (11/44) of MRI scans taken during the acute relapse phase showed T1 enhancing lesions which were identified in 17.5% (27/154) of the MRI scans taken during periods of remission.

Classification results. We trained a ResNet-18 model that can take 5 channels as input for 25 epochs. We created a 5-channel 2D image by concatenating the selected five axial slices, which we used as input data. Group K-fold is a K-fold validation method that prevents multiple images of one patient from being included in some training set and other images from being included in the validation or test dataset. Batch size was 10, loss function was optimized using the Adam optimizer, and the learning rate was set to $5e-4$. The weighted CrossEntropyLoss function was applied to solve the imbalance of the images used. We only used augmented data

	Enrolled patients with MS N = 86	Enrolled patients with NMOSD N = 70	P value
Age, years (SD)	35.0 (9.9)	43.9 (12.6)	< 0.001
Female, n (%)	62 (72.1)	60 (85.7)	0.063
AQP4 antibody, n (%) ^a	0 (0)	66 (94.3)	< 0.001
Total number of MRI scans	199	109	< 0.001
Baseline scans, n (%)	86 (43.2)	70 (64.2)	
Follow-up scans, n (%)	113 (56.8)	39 (35.8)	
Disease duration at the time of MRI, years (SD)	5.8 (5.3)	4.8 (5.4)	0.020
EDSS at the time of MRI, median (IQR)	1.00 (0–2.5)	2.5 (1.5–4.0)	< 0.001
Clinical status at the time of MRI			0.106
In acute relapse, n (%)	44 (22.1)	34 (31.2)	
In remission, n (%)	155 (77.9)	75 (68.8)	
Characteristics of brain MRI in NMOSD ^b			
Normal MRI	N/A	21 (19.3)	
Non-specific brain lesions	N/A	38 (34.9)	
Longitudinal corticospinal tracts lesion, n (%)	N/A	11 (10.1)	
Extensive hemispheric lesion, n (%)	N/A	16 (14.7)	
Periependymal lesion, n (%)	N/A	42 (38.5)	
Cervicomedullary lesion, n (%)	N/A	6 (5.5)	
Barkhof criteria fulfillment, n (%)	164 (82.4)	N/A	
Presence of T1 enhancing MS lesion, n (%) ^c	38/198 (19.2)	N/A	

Table 1. Demographic characteristics of enrolled patients with multiple sclerosis and neuromyelitis optica spectrum disorder. MS, Multiple sclerosis; NMOSD, Neuromyelitis optica spectrum disorder; SD, Standard deviation; AQP4, Aquaporin-4; MRI, Magnetic resonance imaging; IQR, inter-quartile range; N/A, not applicable. ^aMOG antibody was negative for all enrolled patients. ^bA single MRI scan can have multiple findings. ^cOne MRI scan was taken without enhancement.

during model training, not for validation or testing. The test set was used to evaluate the ultimate performance of the model. To minimize the influence on the model's assessment of cases where follow-up scans might not show any major differences or may greatly resemble the baseline scans, the test set consisted only of baseline images of patients without any follow-up images. From the pool of MRI scans of 34 MS patients and 46 NMOSD patients, all of whom had not undergone any follow-up scans, we randomly selected 15 MRI images from each group. Using a different random seed each time, we repeated this process 100 times. The classification results are presented in Table 2. The accuracy of this model in differentiating between NMOSD and MS was 76.1% (95% CI 74.8–77.4) with a sensitivity of 77.3% (95% CI 74.4–80.3) and a specificity of 74.8% (95% CI 72.1–77.5). Positive predictive value (PPV) and negative predictive value (NPV) were 76.9% (95% CI 75.2–78.7) and 78.6% (95% CI 76.7–80.6), with an area under the receiver operating characteristic (ROC) curve of 0.85 (95% CI 0.84–0.86).

Gradient-weighted class activation map (Grad-CAM). We generated a gradient-weighted class activation map (Grad-CAM) to evaluate if the five 2D slices selected from the 3D fluid-attenuated inversion-recovery (FLAIR) images represented lesions that could be used to distinguish between MS and NMOSD. Grad-CAM results are shown in Fig. 1. Areas with white matter lesions are highlighted in red, indicating that our ResNet-18 model generated results by recognizing MS and NMOSD lesions in the images during the classification task.

Discussion

We developed a compact deep learning model with good accuracy and prediction using five axial slices of FLAIR brain MRI for differentiating MS and NMOSD. Further exploration of this model using Grad-CAM showed that white matter lesions were what the model focused on for classification.

Diagnosis of MS can be challenging if patients have atypical clinical presentations. Misdiagnosis of MS could cause patients to undergo hazardous treatment; MS therapies, including interferon beta or fingolimod, can

	Accuracy, %	Sensitivity, %	Specificity, %	PPV	NPV	AUC
Model ^a	76.1 (74.8–77.4)	77.3 (74.4–80.3)	74.8 (72.1–77.5)	76.9 (75.2–78.7)	78.6 (76.7–80.6)	0.85 (0.84–0.86)

Table 2. Classification results of multiple sclerosis and neuromyelitis optica spectrum disorder using the proposed architecture. PPV, positive predictive value; NPV, negative predictive value; AUC, area under the curve. ^aData have been calculated with 100 times experiments and presented with mean and 95% confidence interval.

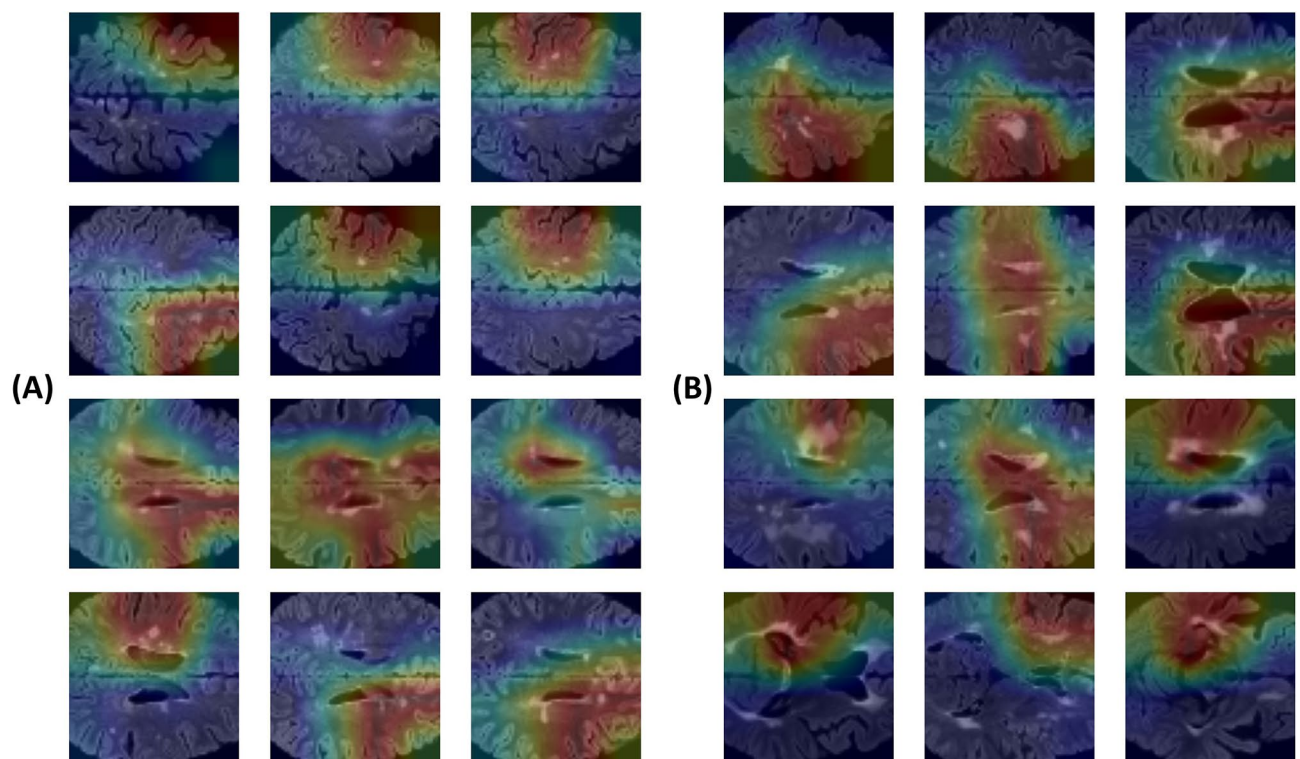


Figure 1. Application of Grad-CAM. (A) Grad-CAM results for MS input data, and (B) Grad-CAM results for NMOSD. The areas marked in red includes white matter lesions in MS and NMOSD. Grad-CAM, Gradient-weighted Class Activation Map; MS, Multiple Sclerosis; NMOSD, Neuromyelitis Optica Spectrum Disorder.

exacerbate NMOSD^{17,18}. Serologic testing, which is the major diagnostic criterion for NMOSD, can help differentiate MS from NMOSD, but there are still limitations in the availability of antibody testing, and seronegative cases exist. Misdiagnosis of MS is common, and revision of the McDonald criteria in 2017 raised concerns about misdiagnosis and emphasized the need for systematic identification of typical MRI features, but exclusion of alternative diagnoses is not standardized^{1,11,19,20}. Characteristics of brain MRI lesions have also been studied to differentiate between MS and NMOSD. However, we showed that 10% of brain MRIs of patients with the onset of NMOSD met the MS MRI criteria, suggesting it may be challenging to distinguish NMOSD from MS based only on brain MRI at onset¹⁵. Previously, brain lesions characteristic of NMOSD were observed in 69% of patients with NMOSD during the disease course¹⁶. Other cross-sectional studies showed lower frequencies of NMOSD-specific brain lesions: 50.9% in chronic phase European patients and 17.7% in chronic phase Chinese patients^{21,22}. This indicates that lesions characteristic of NMOSD could be missed outside the acute phase and that different ethnic population, selection bias, or expert knowledge could affect accurate differentiation of the two disorders^{6,15}. In our study, only 38.7% of MRI scans performed in the chronic phase showed NMOSD-specific brain lesions, suggesting that it might be challenging to distinguish NMOSD based on MRI in our study population.

Machine learning is an alternative approach to differentiating between NMOSD and MS. Efforts have been made to apply machine learning algorithms to differentiate between MS and NMOSD. Multiple modalities, including functional MRI, white matter lesions, gray matter measures, diffusion tensor imaging, cortical thickness, and cognitive/clinical assessment, were used; a high accuracy of 74% to 84% was attained depending on modality, which can improve our understanding of the characteristics of the disease related to the modalities^{13,23}. However, the models used were not fully automated and requiring expert evaluation and selection of the features.

Deep learning models can overcome these obstacles. Only two studies have applied deep learning-based methods to distinguish MS from NMOSD. One study reported 81.3% accuracy of differentiation between MS and NMOSD using hierarchical multimodal fusion models that integrated FLAIR and diffusion tensor imaging (DTI) sequences²⁴; the other showed 71.1% accuracy using CNN integrated brain MRI and clinical data¹⁴. Our deep learning model used only five axial slices of FLAIR MRI data, and showed comparable accuracy (76.1%) with good sensitivity and specificity (77.3%, and 74.8%, respectively).

The deep learning model we used is the residual neural network (ResNet), which is a neural network widely used in the medical field^{25–28}. ResNet architecture solves the issue of a relatively limited training dataset and enhances image classification performance by expanding the network's depth. Skip connections are used by ResNet to alleviate degradation issues. Empirical evidence suggests that the training cost of 3D Convolutional neural networks (CNNs) is significantly greater than that of 2D CNNs as more trainable parameters necessitate longer training durations and more training data. The limited dataset size used in this study may potentially have impacted the function of 3D CNNs. Consequently, we employed ResNet 2D CNNs in our study. Data augmentation is a technique for expanding the amount and quality of training datasets to improve the performance of deep learning models. Additionally, data augmentation involves adding missing data points to the initial training data²⁹. Using this approach, we attempted to minimize potential issues associated with using the limited amount of FLAIR MRI data. This approach has been used in other studies; a classification task was performed with ResNet and data augmentation with flipping was found to enhance accuracy^{30,31}, while a noisy augmented dataset offered superior classification accuracy on ResNet compared to the original dataset³².

The complexity of the learning process makes it challenging to interpret deep learning models³³. Grad-CAM method can provide insight into how deep learning models classify images by facilitating localization of features that the deep learning model focuses on using a heatmap^{33,34}; a deep learning model may distinguish between images in ways that are distinct from how humans do³⁵. In this study, Grad-CAM revealed that the model focused on white matter lesions to differentiate between MS and NMOSD (Fig. 1). Unknown features of two diseases other than white matter lesions were not recognized with Grad-CAM; white matter lesions therefore appear to be an appropriate area for classification. Further deep learning models with large scale image data from MS and NMOSD could help discover new imaging characteristics.

This study has several limitations. First, this study was conducted with a relatively small number of MRIs in a single center without external validation, which limits generalization of our findings. Second, our model was trained for binary classification, and brain MRIs of healthy subjects were not included in this study. This could be a significant barrier when implementing this model in clinical settings. Third, the clinical state of the disease when MRI scans were performed was not controlled; 68.8% of MRI scans were taken in a chronic remission state. However, given that it may be more challenging to differentiate NMOSD in the chronic phase using MRI data than NMOSD in the acute phase, our findings suggest that this model is useful. Further investigations with extensive data are required to develop a fully automated deep learning model for the diagnosis of CNS demyelinating diseases.

In conclusion, we developed a compact deep learning model based on FLAIR brain MRI data with the ability to differentiate MS from NMOSD. We showed that this model, using the Grad-CAM approach, differentiated between MS and NMOSD based on white matter lesions. This compact deep learning model may aid in the differential diagnosis of MS from NMOSD in clinical practice.

Methods

Patients. We prospectively evaluated patients who visited the neurology outpatient clinic of Samsung Medical Center (Seoul, Korea) between May 2016 and May 2020. Patients were enrolled if they had MS or NMOSD, and their diagnosis was performed by two experienced neurologists according to the 2017 McDonald criteria or the international consensus diagnostic criteria for NMOSD, respectively^{1,2}. We collected brain MRIs during clinical follow-up; standardized T2-weighted, three-dimensional T1-weighted turbo field echo, and three-dimensional fluid-attenuated inversion recovery images were acquired using a 3.0-T MRI scanner (Philips 3.0 T

Achieva, Philips Healthcare, Andover, MA, USA) as described previously³⁶. Patients were excluded from the study if (a) AQP4 and MOG antibodies were not assessed, (b) they declined to participate in the study, and (c) they had a history of brain surgery or medical disorders, including cerebral infarction, intracranial hemorrhage, brain tumor or head trauma as these can alter brain MRI findings. We also collected demographic characteristics of the enrolled patients, including gender, age, and seropositivity for AQP4 and MOG antibodies.

The study and all experimental protocols were approved by the institutional review board (IRB) of the Samsung Medical Center; all participants provided written informed consent prior to the commencement of the study, and all methods were performed in accordance with the relevant guidelines and regulations.

Image preprocessing. Preprocessing is a set of operations performed on an image to improve its quality and make statistical analysis more repeatable and comparable. Image registration is a critical step in various biomedical imaging applications. It provides the ability to align one image with another geometrically and is a prerequisite for all imaging applications that compare datasets across subjects, imaging modalities, or time³⁷. We registered FLAIR images to T1 images. This was done using FMRIB (Functional Magnetic Resonance Imaging of the Brain)'s Linear Image Registration Tool (FLIRT). The overall geometry of the brain is unlikely to be altered for scans from the same individual, but each scan may have experienced a translation and/or rotation in space. We employed rigid-body transformation with 12 degrees of freedom to correct for this. We used Freesurfer 6.0 to resample the FLAIR image to 256 size and correct intensity non-uniformity³⁸. T1 images were converted to Montreal Neurological Institute (MNI) standard space using FMRIB's Nonlinear Image Registration Tool (FNIRT), and coefficient maps were obtained in this process. FLAIR images were converted to MNI standard space using FSL's applywarp function, which applies the FNIRT's coefficient map to other images. To obtain only the brain part without the background, we cropped the FLAIR image to 128 size.

A total of five axial slices were chosen at 20 slice intervals before and after to distinguish multiple sclerosis from NMOSD based on the position of the lateral ventricle where lesions are present in both disorders but the morphology of the lesions differs (Fig. 2)¹⁰. The five axial slice positions were the most similar positions presented on the report of Matthews and colleagues⁹, representing the cortical area, deep white matter area,

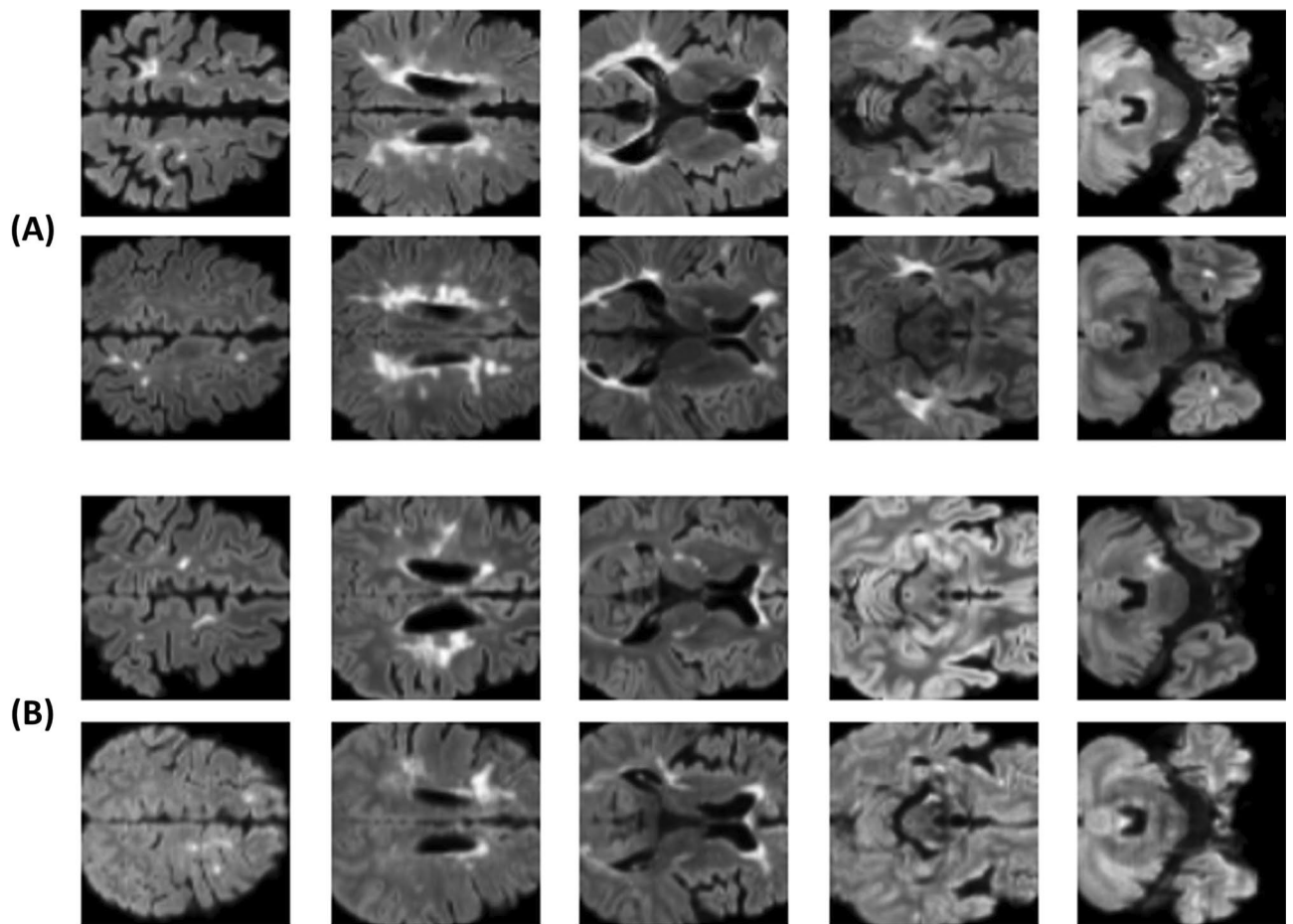


Figure 2. Example of 2D FLAIR image input data for use in the classification model. (A) Images from two patients with multiple sclerosis, (B) images from two patients with neuromyelitis optica spectrum disorder. One slice is one channel, so the five images in a row were merged into one input image with five channels.

lateral ventricle, basal ganglia, and brainstem/cerebellum. We replaced one slice with one channel, resulting in a five-channel input image.

Convolutional neural networks. CNNs is a deep learning method that trains several layers. It is used for a variety of computer vision applications and is very efficient^{39–41}. In general, a CNN consists of three main neural layers: convolutional layers, pooling layers, and fully connected layers. Convolutional layers are at the core of a CNN. Convolution is a linear process that, like a conventional neural network, multiplies a set of weights with the input in the context of a convolutional neural network. Multiplication is done between an input array and a two-dimensional array of weights, known as a filter or a kernel, because the approach was designed for two-dimensional input data. A single value is produced by multiplying the filter by the input array once. A two-dimensional array of output values representing an input filter is produced when the filter is applied to the input array more than once. The two-dimensional output array from this operation is known as a feature map. Once a feature map has been generated, each value is passed through a nonlinearity. The function of the pooling layer is to reduce the dimensions by pooling feature maps. It also collects and enhances the features of the extracted image. A fully connected layer is used in a classification task, and a likelihood function is used to calculate the likelihood probability of each image class from the fully connected layer. The most probable labels serve as classifiers throughout the CNN and are output as classification results.

Data augmentation. High-quality, abundant data is critical in the development of deep learning models. A deficit of training data can lead to overfitting⁴². The classification problem addressed in this paper lacks sufficient data to provide a deep learning architecture. Therefore, we performed data augmentation based on the training set using the following two methods to achieve the desired accuracy. The first data augmentation method we used was the RandomHorizontalFlip. RandomHorizontalFlip is a type of image data augmentation that flips the input image horizontally with a given probability. The second data augmentation method we used is RandomNoise⁴³. RandomNoise is a simple form of data augmentation that adds noise sampled from a normal, random distribution. By training a neural network on noisy data, robust neural networks that proficiently generalize, even on noisy images, can be generated.

Model architecture. We used a model based on the ResNet CNN model⁴⁴. There are several types of ResNet, such as ResNet-18, ResNet-50, and ResNet-101. In ResNet-n, n is the number of layers in the network, and as n increases, the number of computations increases, as well as the performance of the network. We used ResNet-18 with some changes; ResNet18 is a CNN model with a 72-layer architecture and 18 deep layers. ResNet18 consists of one 7×7 convolutional layer, two pool layers, eight residual units, and one fully connected layer. Each of the residual units contains two 3×3 convolutional layers. Here, we changed the input image of ResNet-18 to five channels and the output class to two types. Figure 3 shows the modified ResNet architecture used in this study to differentiate between MS and NMOSD.

Gradient-weighted class activation map (Grad-CAM). Grad-CAM is a generalization of the class activation map (CAM) that finds weights through gradients as follows^{34,45}:

$$\alpha_k^c = \frac{1}{z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k}$$

In the final convolutional layer, we allowed the gradients of any target concept score (logits for any class of interest) to flow. Specific aspects in the image for predicting that concept could then be highlighted on a coarse localization map by computing a significance score based on the gradients. To express this more technically, we computed the gradient of the class C logits concerning the activation maps of the final convolutional layer. Then we averaged the gradients over each feature map to determine a significance score as expressed below:

$$L_{\text{Grad-CAM}}^c = \text{ReLU} \left(\sum_k \alpha_k^c A^k \right)$$

where c is the class of interest, k is the index of the activation map in the final convolutional layer, y^c is the score for class c before softmax, and A^k is the feature map of the k-th channel of the last CNN layer. The alpha value indicates the significance of feature map k for the target class c. The values are then added together after multiplying each activation map by its significance score. ReLU nonlinearity is also used in the summation to take into account only those pixels that positively affect the score of the class of interest.

Statistical analysis. Clinical characteristics of the enrolled patients are presented with appropriate summary statistics. Continuous data are shown as means with standard deviations or medians with inter-quartile ranges (IQRs). Categorical variables are presented as absolute and relative frequencies. We compared demographic findings between the two groups (MS versus NMOSD) using the Chi-square test or Fisher's exact test for categorical variables. Student's t-tests or Mann–Whitney U tests were used to compare continuous variables. The performance of our model was evaluated using appropriate classification metrics, namely accuracy, sensitivity, specificity, PPV, NPV, and area under the ROC curve. The results of 100 experiments are presented with means and 95% confidence intervals (CIs). All statistical analyses were performed using SPSS for Windows version

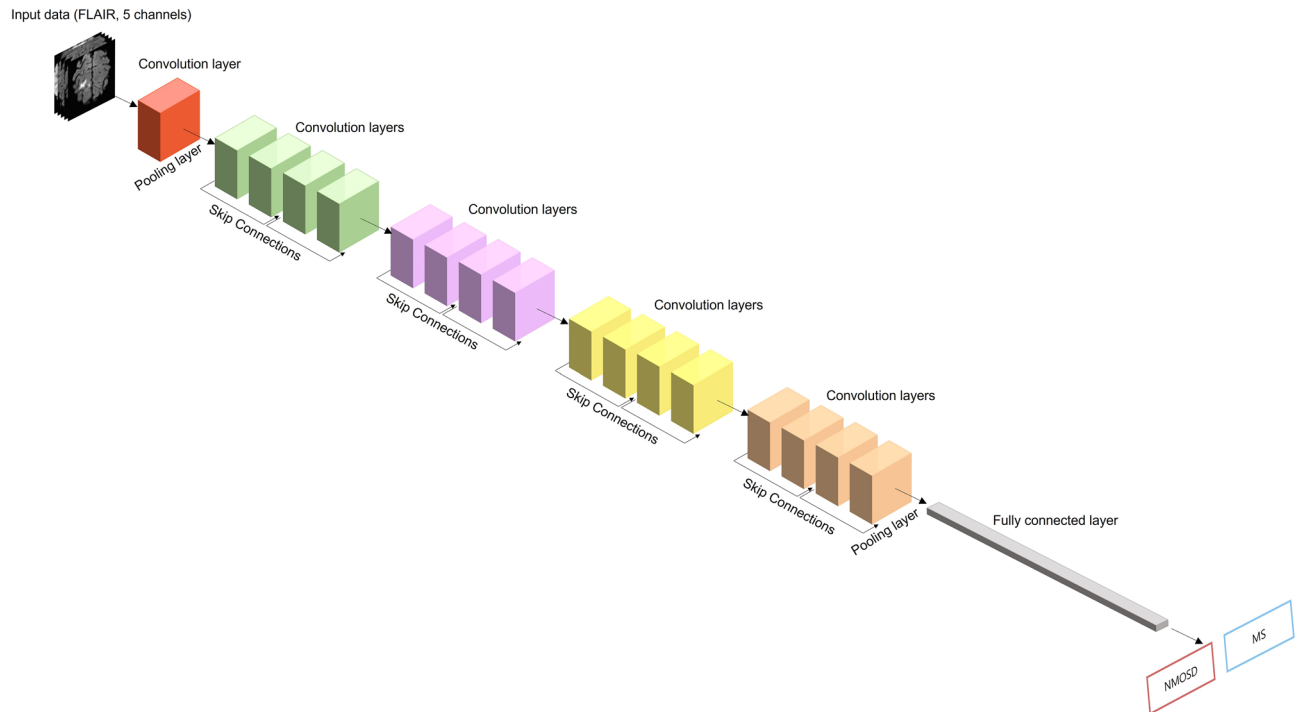


Figure 3. The architecture used to distinguish neuromyelitis optica spectrum disorder from multiple sclerosis. This architecture is based on a residual neural network. The output class of ResNet-18 was changed to 2 and the input data was expanded to 5 channels.

20.0 (IBM, Armonk, NY, USA) or R software version 4.2.1. Statistical significance was defined as a two-tailed p -value < 0.05 .

Data availability

The datasets used and/or analyzed during the current study can be available from the corresponding author on reasonable request.

Received: 23 January 2023; Accepted: 6 July 2023

Published online: 19 July 2023

References

- Thompson, A. J. *et al.* Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol.* **17**, 162–173. [https://doi.org/10.1016/S1474-4422\(17\)30470-2](https://doi.org/10.1016/S1474-4422(17)30470-2) (2018).
- Wingerchuk, D. M. *et al.* International consensus diagnostic criteria for neuromyelitis optica spectrum disorders. *Neurology* **85**, 177–189. <https://doi.org/10.1212/WNL.0000000000001729> (2015).
- Jarius, S. *et al.* Neuromyelitis optica. *Nat. Rev. Dis. Primers* **6**, 85. <https://doi.org/10.1038/s41572-020-0214-9> (2020).
- Waters, P. J. *et al.* Serologic diagnosis of NMO: A multicenter comparison of aquaporin-4-IgG assays. *Neurology* **78**, 665–671. <https://doi.org/10.1212/WNL.0b013e318248dec1> (2012) (**discussion 669**).
- Marchionatti, A., Woodhall, M., Waters, P. J. & Sato, D. K. Detection of MOG-IgG by cell-based assay: Moving from discovery to clinical practice. *Neurol. Sci.* **42**, 73–80. <https://doi.org/10.1007/s10072-020-04828-1> (2021).
- Cortese, R. *et al.* Differentiating multiple sclerosis from AQP4-neuromyelitis optica spectrum disorder and MOG-antibody disease with imaging. *Neurology* <https://doi.org/10.1212/WNL.0000000000201465> (2022).
- Matsuoka, T. *et al.* Heterogeneity of aquaporin-4 autoimmunity and spinal cord lesions in multiple sclerosis in Japanese. *Brain* **130**, 1206–1223. <https://doi.org/10.1093/brain/awm027> (2007).
- Pisani, F. *et al.* Aquaporin-4 autoantibodies in neuromyelitis optica: AQP4 isoform-dependent sensitivity and specificity. *PLoS ONE* **8**, e79185. <https://doi.org/10.1371/journal.pone.0079185> (2013).
- Matthews, L. *et al.* Distinction of seropositive NMO spectrum disorder and MS brain lesion distribution. *Neurology* **80**, 1330–1337. <https://doi.org/10.1212/WNL.0b013e3182887957> (2013).
- Kim, H. J. *et al.* MRI characteristics of neuromyelitis optica spectrum disorder: An international update. *Neurology* **84**, 1165–1173. <https://doi.org/10.1212/WNL.0000000000001367> (2015).
- Solomon, A. J., Naismith, R. T. & Cross, A. H. Misdiagnosis of multiple sclerosis: Impact of the 2017 McDonald criteria on clinical practice. *Neurology* **92**, 26–33. <https://doi.org/10.1212/WNL.0000000000006583> (2019).
- Patel, U. K. *et al.* Artificial intelligence as an emerging technology in the current care of neurological disorders. *J. Neurol.* **268**, 1623–1642. <https://doi.org/10.1007/s00415-019-09518-3> (2021).
- Eshaghi, A. *et al.* Gray matter MRI differentiates neuromyelitis optica from multiple sclerosis using random forest. *Neurology* **87**, 2463–2470. <https://doi.org/10.1212/WNL.0000000000003395> (2016).
- Kim, H. *et al.* Deep learning-based method to differentiate neuromyelitis optica spectrum disorder from multiple sclerosis. *Front. Neurol.* **11**, 599042. <https://doi.org/10.3389/fneur.2020.599042> (2020).
- Huh, S. Y. *et al.* The usefulness of brain MRI at onset in the differentiation of multiple sclerosis and seropositive neuromyelitis optica spectrum disorders. *Mult. Scler.* **20**, 695–704. <https://doi.org/10.1177/1352458513506953> (2014).

16. Kim, W. *et al.* Characteristic brain magnetic resonance imaging abnormalities in central nervous system aquaporin-4 autoimmunity. *Mult. Scler.* **16**, 1229–1236. <https://doi.org/10.1177/1352458510376640> (2010).
17. Min, J. H., Kim, B. J. & Lee, K. H. Development of extensive brain lesions following fingolimod (FTY720) treatment in a patient with neuromyelitis optica spectrum disorder. *Mult. Scler.* **18**, 113–115. <https://doi.org/10.1177/1352458511431973> (2012).
18. Palace, J., Leite, M. I., Nairne, A. & Vincent, A. Interferon Beta treatment in neuromyelitis optica: Increase in relapses and aquaporin 4 antibody titers. *Arch. Neurol.* **67**, 1016–1017. <https://doi.org/10.1001/archneurol.2010.188> (2010).
19. Kaisey, M., Solomon, A. J., Luu, M., Giesser, B. S. & Sciotte, N. L. Incidence of multiple sclerosis misdiagnosis in referrals to two academic centers. *Mult. Scler. Relat. Disord.* **30**, 51–56. <https://doi.org/10.1016/j.msard.2019.01.048> (2019).
20. Geraldes, R. *et al.* The current role of MRI in differentiating multiple sclerosis from its imaging mimics. *Nat. Rev. Neurol.* **14**, 199–213. <https://doi.org/10.1038/nrneurol.2018.14> (2018).
21. Cacciaguerra, L. *et al.* Brain and cord imaging features in neuromyelitis optica spectrum disorders. *Ann. Neurol.* **85**, 371–384. <https://doi.org/10.1002/ana.25411> (2019).
22. Cao, G. *et al.* Brain MRI characteristics in neuromyelitis optica spectrum disorders: A large multi-center retrospective study in China. *Mult. Scler. Relat. Disord.* **46**, 102475. <https://doi.org/10.1016/j.msard.2020.102475> (2020).
23. Eshaghi, A. *et al.* Classification algorithms with multi-modal data fusion could accurately distinguish neuromyelitis optica from multiple sclerosis. *Neuroimage Clin.* **7**, 306–314. <https://doi.org/10.1016/j.nicl.2015.01.001> (2015).
24. Yoo, Y. *et al.* *Medical Image Computing and Computer Assisted Intervention—MICCAI 2017* 480–488 (Springer International Publishing, 2017).
25. Ayyachamy, S., Alex, V., Khened, M. & Krishnamurthi, G. *Medical Image Retrieval Using Resnet-18*. vol. 10954 MI (SPIE, 2019).
26. Xu, H., Liu, Y., Zeng, X., Wang, L. & Wang, Z. In *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. 2153–2156.
27. Yuanyuan, P., Weifang, Z., Feng, C., Daoman, X. & Xinjian, C. In *Proc.SPIE*. 1131321.
28. Uyulan, C. *et al.* A class activation map-based interpretable transfer learning model for automated detection of ADHD from fMRI data. *Clin. EEG Neurosci.* <https://doi.org/10.1177/15500594221122699> (2022).
29. Shorten, C. & Khoshgoftaar, T. M. A survey on image data augmentation for deep learning. *J. Big Data* **6**, 60. <https://doi.org/10.1186/s40537-019-0197-0> (2019).
30. Valliani, A. & Soni, A. In *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*. 615–615.
31. Ghosal, P. *et al.* In *2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP)*. 1–6 (IEEE).
32. Pandian, J. A., Geetharamani, G. & Annette, B. In *2019 IEEE 9th international conference on advanced computing (IACC)*. 199–204 (IEEE).
33. Zhang, Y. *et al.* Grad-CAM helps interpret the deep learning models trained to classify multiple sclerosis types using clinical brain magnetic resonance imaging. *J. Neurosci. Methods* **353**, 109098. <https://doi.org/10.1016/j.jneumeth.2021.109098> (2021).
34. Selvaraju, R. R. *et al.* Grad-CAM: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 618–626 (2017).
35. Iizuka, T., Fukasawa, M. & Kameyama, M. Deep-learning-based imaging-classification identified cingulate island sign in dementia with Lewy bodies. *Sci. Rep.* **9**, 8944. <https://doi.org/10.1038/s41598-019-45415-5> (2019).
36. Seok, J. M. *et al.* Association of subcortical structural shapes with fatigue in neuromyelitis optica spectrum disorder. *Sci. Rep.* **12**, 1579. <https://doi.org/10.1038/s41598-022-05531-1> (2022).
37. Toga, A. W. & Thompson, P. M. The role of image registration in brain mapping. *Image Vis. Comput.* **19**, 3–24. [https://doi.org/10.1016/S0262-8856\(00\)00055-X](https://doi.org/10.1016/S0262-8856(00)00055-X) (2001).
38. Fischl, B. FreeSurfer. *Neuroimage* **62**, 774–781. <https://doi.org/10.1016/j.neuroimage.2012.01.021> (2012).
39. Long, J., Shelhamer, E. & Darrell, T. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 3431–3440.
40. Li, Q. *et al.* In *2014 13th International Conference on Control Automation Robotics & Vision (ICARCV)*. 844–848 (IEEE).
41. Gidaris, S. & Komodakis, N. In *Proceedings of the IEEE International Conference on Computer Vision*. 1134–1142.
42. Ying, X. An overview of overfitting and its solutions. *J. Phys. Conf. Ser.* **1168**, 022022 (2019).
43. Perez-Garcia, F., Sparks, R. & Ourselin, S. TorchIO: A Python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning. *Comput. Methods Programs Biomed.* **208**, 106236. <https://doi.org/10.1016/j.cmpb.2021.106236> (2021).
44. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778 (2016).
45. Zhou, B., Khosla, A., Lapedriza, A., Oliva, A. & Torralba, A. Learning deep features for discriminative localization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2921–2929 (2016).

Author contributions

J.M.S.: conceptualization, data analysis, manuscript writing. W.C.: data analysis, manuscript writing. Y.H.C.: data collection, data analysis. H.J.: data collection, data analysis. S.T.K.: data collection. J-K.S.: conceptualization, data analysis, manuscript editing. J-H.M.: conceptualization, data collection, data analysis, manuscript editing.

Funding

This research was supported by a grant from the National Research Foundation of Korea (NRF) funded by the Korean government (MSIT) (2021R1F1A1049347 to JHM).

Competing interests

J.M.S. reports no relevant disclosures. W.C. reports no relevant disclosures. Y.H.C. reports no relevant disclosures. H.J. reports no relevant disclosures. S.T.K. reports no relevant disclosures. J-K.S. reports no relevant disclosures. J-H.M. has received funding and research support from the National Research Foundation of Korea and SMC Research and Development Grants and has lectured, consulted, and received honoraria from Bayer Schering Pharma, Merck, Biogen Idec, Sanofi, UCB, Samsung Bioepis, Mitsubishi Tanabe, Celltrion, Roche, and Janssen.

Additional information

Correspondence and requests for materials should be addressed to J.-K.S. or J.-H.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023