Diagnostic effectiveness of deep learning-based MRI in predicting multiple sclerosis: A meta-analysis

Tareef S. Daqqaq, MBBS, Facharzt, *Ayman S. Alhasan,* MBBS, DES, *Hadeel A. Ghunaim,* MBBS.

ABSTRACT

الأهداف: يمكن أن يتأثر الدماغ والحبل الشوكي، اللذان يشكلان الجهاز العصبي المركزي (CNS)، بمرض التهابي يعرف باسم التصلب اللويحي (MS). يمكن للشبكات العصبية التلافيفية (CNN)، وهي من طرق الذكاء الاصطناعي، والتي تستطيع اكتشاف الالتهاب مبكرًا من خلال تعلم الأنماط على صورة الرنين المغناطيسي للدماغ(MRI). تم إجراء هذه الدراسة لاستقصاء الأداء التشخيصي للتصوير بالرنين المغناطيسي المعتمد على الشبكات العصبية التلافيفي CNN، في تحديد وتصنيف وتقسيم آفات مرض التصلب اللويحي.

المنهجية: تم استخدام محركات البحث Web of Science، PubMed Embase، Cochrane Library، CINAHL، Google Scholar لاسترداد الأوراق التي تشير إلى استخدام التصوير بالرنين المغناطيسي المعتمد على الشبكات العصبية التلافيفي (CNN)، في تشخيص مرض التصلب اللويحي. تم تقييم الدقة والنوعية والحساسية ومعامل (DSC) في هذه الدراسة.

النتائج: في المجموع، تم تحديد 2174 دراسة و15 مقالة حققت معايير الاشتمال. قدمت 2D-3D CNN دقة عالية (98.81, 95% CI: 98.50–99.13)، وحساسية (98.76, 95% CI: 98.42–99.10)، وخصوصية (98.67, 95% CI: 98.22–99.12) في تحديد آفات مرض التصلب العصبي اللويحي. فيما يتعلق بالتصنيف، كان معدل الدقة الإجمالي مرتفعًا بشكل ملحوظ (91.38, 95% CI: 83.23– 99.54). أظهر DSC معدل 63.78 (95% CI: 58.29–69.27) أن التصوير بالرنين المغناطيسي المستند إلى 2D-3D CNN أداء عالٍ في تجزئة آفات التصلب اللويحي. وأظهر تحليل الحساسية أن النتائج متسقة، مما يشير إلى أن هذه الدراسة قوية.

الخلاصة: كشف هذا التحليل التلوي أن التصوير بالرنين المغناطيسي المعتمد على 2D-3D CNN هو نظام ذكاء اصطناعي يتمتع بأداء تشخيصي عالي ويمكنه التنبؤ بالمرض بسرعة وفعالية

Objectives: The brain and spinal cord, constituting the central nervous system (CNS), could be impacted by an inflammatory disease known as multiple sclerosis (MS). The convolutional neural networks (CNN), a machine learning method, can detect lesions early by learning patterns on brain magnetic resonance image (MRI). We performed this study to investigate the diagnostic performance of CNN based MRI in the identification, classification, and segmentation of MS lesions.

Methods: PubMed, Web of Science, Embase, the Cochrane Library, CINAHL, and Google Scholar were used to retrieve papers reporting the use of CNN based MRI in MS diagnosis. The accuracy, the specificity, the sensitivity, and the Dice Similarity Coefficient (DSC) were evaluated in this study.

Results: In total, 2174 studies were identified and only 15 articles met the inclusion criteria. The 2D-3D CNN presented a high accuracy (98.81, 95% CI: 98.50–99.13), sensitivity (98.76, 95% CI: 98.42–99.10), and specificity (98.67, 95% CI: 98.22–99.12) in the identification of MS lesions. Regarding classification, the overall accuracy rate was significantly high (91.38, 95% CI: 83.23–99.54). A DSC rate of 63.78 (95% CI: 58.29–69.27) showed that 2D-3D CNN-based MRI performed highly in the segmentation of MS lesions. Sensitivity analysis showed that the results are consistent, indicating that this study is robust.

Conclusion: This metanalysis revealed that 2D-3D CNN based MRI is an automated system that has high diagnostic performance and can promptly and effectively predict the disease.

Neurosciences 2020; Vol. 25 (3): 77-89
doi: 10.17712/nsj.2024.2.20230103

From the Department of Internal Medicine (Daqqaq, Alhasan, Ghunaim),College of Medicine, Taibah University, Madinah, and from Department of Radiology (Daqqaq), Prince Mohammed Bin Abdulaziz Hospital, Ministry of National Guard Health Affairs, and from the Department of Radiology (Alhasan), King Faisal Specialist Hospital and Research Center, Madinah, Kingdom of Saudi Arabia.

Received 15th October 2023. Accepted 6th January 2024.

Address correspondence and reprint request to: Dr. Tareef S. Daqqaq, Department of Internal Medicine, College of Medicine, Taibah University, Madinah, Kingdom of Saudi Arabia. E-mail: tdaqqaq@taibahu.edu.sa
ORCID ID: https://orcid.org/0000-0002-1479-0897

Multiple sclerosis (MS) is an inflammatory neurological condition that affects the central nervous system (CNS). Specifically in the brain's white

OPEN ACCESS

matter, it causes demyelination and inflammation of the nerves. As a result, it can slow down or block messages between the brain and body. The World Health Organization (WHO) revealed that MS affects 2.8 million people globally, and its prevalence is increasing every year.[1] North America, Europe, and Australia have most of the MS patients.[2] Fatigue, trouble walking, stiffness, weakness, vision issues, vertigo, cognitive changes, emotional changes, sadness, and more are all typical MS symptoms.[3,4] To date, the etiology of MS remains unclear.[5] The MS is thought to be caused by a confluence of hereditary and environmental factors.[6] Geographical location, vitamin D insufficiency, obesity, and smoking are examples of environmental factors that may be related to MS. The early detection of MS presents long-term benefits and could help researchers to find the best clinical strategy of this condition.[7] There are currently no signs, physical observations, or laboratory testing that can alone indicate if you have MS. There are several methods used to assess if you match the recognized standards for an MS diagnosis and to rule out other potential causes of the symptoms you are presently exhibiting. A thorough medical history, a neurologic examination, and numerous diagnostics, such as magnetic resonance imaging (MRI), spinal fluid analysis, and blood testing, are some of these measures.[8] Currently, the standard non-invasive diagnostic modality of MS uses MRI to visualize the lesions and presents a major role in controlling prognosis and development of the disease.[9] The use of quantitative MRI techniques improves understanding of the extent of tissue damage and disease. These methods consist of: (1) MR spectroscopy, a non-invasive technique for examining the biochemical changes in MS;[10] (2) magnetization transfer imaging which offers improved sensitivity and specificity for MS studies;[11] (3) diffusion weighted imaging (DWI) and diffusion tensor imaging (DTI) which are quantitative MRI techniques, providing information on size, integrity, geometry, and orientation of tissue fibers;[12] (4) dynamic contrast enhanced MRI which enables quantification of blood brain barrier disruption;[13] and (5) dynamic susceptibility contrast MRI, which by injecting contrast agent into the patient produces quantitative maps of cerebral blood flow, cerebral blood volume, and temporal metrics like mean transit time.[13] To date, the accurate diagnosis of MS lesions presents some inconvenient. It was showed that MRI are difficult and time-consuming modalities in the diagnosis of MS because it is difficult to manually detect most of the lesions, especially within the grey matter.[14] Moreover, interobserver variability can lead to inaccurate results, and the inability to compare studies from several modalities is a significant drawback. Also, methods that include human interplay may be characterized by substantial inter observer variability that may hamper the quality of the final results. Thus, new modalities were proposed to correctly detect MS lesions. Recently, deep learning (DL) tools, using artificial intelligence (AI), have developed for the diagnosis of various diseases, attracting many physicians' attention.[15,16] Different DL techniques using MRI were proposed for the diagnosis of MS. The principal advantage of DL methods is their capacity to deduct intrinsic image representation in MRI data.[17] Furthermore, DL does not need any manual guidance of the characteristic extraction step.[18]

Since 2016, research on the use of DL architectures and MRI data for the diagnosis of MS have been conducted. Identification, segmentation, and classification of MS lesions were investigated by DL models. Conventional neural networks (CNN) are one of the most widely employed architectures in MS diagnosis.[14] It learns characteristics of lesions using multinomial logistic regression to improve the diagnosis of MS.[19] The majority of physicians used 2D- and 3D-CNN architectures for classification and segmentation of MRI techniques. These networks' capacity to reuse weights and lower parameter counts make them more compatible with 2D and 3D images.[14] Given that 3D images contain a lot more information than 2D images, it makes some sense that the 3D CNN will perform better than the 2D version. The CNN designs and/or training/testing dataset variances may be the root cause of the underlying variations. A 3D CNN requires far more processing power for training and inference than a 2D CNN.

The objective of this metaanalysis is to assess the effectiveness of MRI based 2D-3D CNN architectures on the diagnosis of MS.

Methods. *Resources and search techniques.* We followed the Preferred Reporting Items for Systematic Reviews and MetaAnalyses (PRISMA) standards to conduct this study.[20] The following databases were used to find relevant papers published from January 2010 until December 2022: Web of Science, PubMed, CINAHL, Google Scholar, Embase, and the Cochrane Library. Two independent reviewers performed a

Disclosure. Authors have no conflict of interests, and the work was not supported or funded by any drug company.

systematic search using the following terms "multiple sclerosis" AND "magnetic resonance imaging" OR "MRI" AND "machine learning" OR "artificial intelligence" OR "deep learning" OR "convolutional neural networks".

Selection criteria. After suppression of duplicates, title and abstract checks were done on pertinent papers. Papers were included if they reported the use of 2D- or 3D-CNN for the identification, classification, or segmentation of MS lesions. They were then fully read to ensure eligibility.

Study inclusion criteria were: (1) Papers reporting MS lesions; (2) identification, classification, or segmentation of MS lesions using a CNN method; (3) use of 2D- or 3D-CNN architecture; (4) use of MRI as neuroimaging modality; (5) original research papers; and (6) articles reporting sufficient information about the performance of CNN.

Study exclusion criteria were: (1) Papers written in languages other than English; (2) letters, comments, opinions, guidelines, protocols, and review papers; (3) use of other architectures of CNN (4D-CNN Models, DeepSCAN); (4) overlapping study groups and duplicate publications; (5) studies with scant information on the results.

Data extraction. Two independent authors retrieved information from the eligible articles following the inclusion and exclusion criteria, and information were collected on a standardized data sheet that included: (1) article, (2) country, (3) dataset, (4) sample size, (5) diagnosis application, (6) neuroimaging modalities, (7) deep learning method, (8) deep learning architecture, and (9) performance.

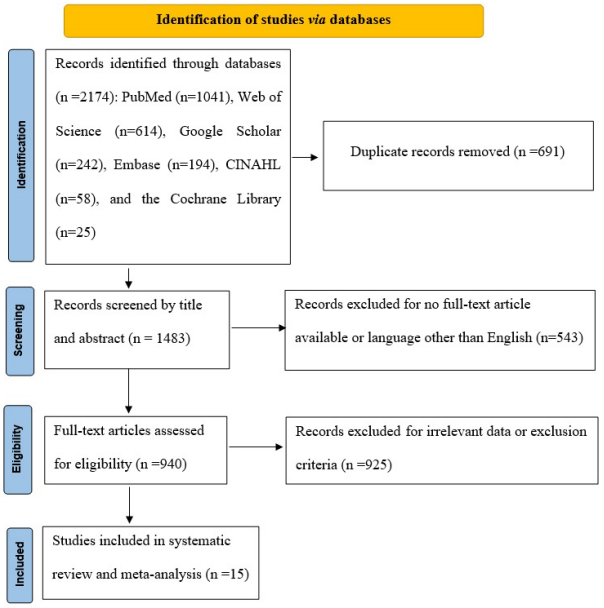

Figure 1 - PRISMA study flowchart.

Study Quality Assessment. The methodologic quality of the included studies was evaluated independently, by 2 authors, using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool, which includes four criteria that judge bias and applicability: "patient selection", "index test", "reference standard", and "flow and timing".[21] Each is assessed in terms of risk of bias, and the first 3 domains were also assessed with respect to applicability. Each item is answered with "yes," "no," or "unclear." The answer of "yes" means low risk of bias, whereas "no" or "unclear" means the opposite. Consensus was used to settle disagreements, while arbitration with a third reviewer was an option if necessary. RevMan Version 5.4 (Cochrane Collaboration, Oxford, United Kingdom) was used to visualize the quality assessment results.

Outcome measures. Accuracy: It measures the ability of 2D- or 3D-CNN to detect MS when it is present and detect the absence of MS when it is absent.

Sensitivity. It refers to 2D- or 3D-CNN's ability to designate an individual with MS as positive.

Specificity. It refers to 2D- or 3D-CNN's ability to correctly classify an individual without MS disease.

Dice Similarity Coefficient (DSC): It is a spatial overlap index and a reproducibility validation metric that measures the similarity between two sets of binary segmentation results.

These measures were used for the :(i) identification of MS from healthy controls, (ii) classification of MS lesions from other brain lesions, and (iii) segmentation of images produced by MRI for measuring and visualizing the brain's anatomical structures, for analyzing brain changes and for delineating MS lesions.

Statistical analysis. Accuracy, sensitivity, specificity, and dice similarity coefficient measures were pooled from the included studies. Statistical analyses were conducted by RevMan Version 5.4 (Cochrane Collaboration, Oxford, United Kingdom). A p-value <0.05 was considered significant. Heterogeneity was assessed by the Cochrane chi-squared test. A p-value <0.05 confirms the presence of heterogeneity. In order to assess the influence of heterogeneity on the results, we calculated I^2 values; I2 values ≥50% and p<0.05 indicated an important level of heterogeneity. If I^2<50% and p>0.05, we used a fixed effects design; if not, a random effects model was adopted.[22] We also performed subgroup and sensitivity analysis to identify the cause of heterogeneity. To assess publication bias, a visual examination of the symmetry in funnel plots was used. This second point was supported by Egger's test using the SPSS V25 statistical package.

Table 1 - Features of included studies

Article	Country	Dataset	Sample size	Diagnosis	Application	Deep learning architecture	Performance
Afzal el al, 2018[15]	Australia	John Hunter Hospital's Dataset	21	-11 converted to MS -10 did not convert to MS	Classification	2D-CNN	Accuracy
Afzal el al, 2021[16]	Australia	ISBI and MICCAI datasets	19	127 scans of MS	Segmentation	2D-CNN	-DSC -Sensitivity -Precision
Alijamaat et al, 2020[17]	Iran	Laboratory of eHealth of the University of Cyprus	58	38 MS patients 20 healthy individuals	Identification	2D-CNN	-Accuracy -Precision -Sensitivity -Specificity
Aslani et al., 2019[18]	Italy	-Private dataset -ISBI 2015 longitudinal dataset	51	-37 patients from private dataset -14 patients from ISBI 2015 longitudinal dataset	Segmentation	2D-CNN	DSC
Aslani et al, 2019[19]	Italy	ISBI 2015 Longitudinal MS Lesion Segmentation	19	MS	Segmentation	2D-CNN	-DSC -Lesion-wise true-positive -Lesion-wise false-positive
Coronado et al, 2020[20]	USA	CombiRx	1,006	Relapsing–remitting MS	Segmentation	3D-CNN	-DSC -Lesion-wise true-positive -Lesion-wise false-positive
Eitel et al, 2019[21]	Germany	Clinical	147	76 MS patients 71 healthy patients	Classification	3D-CNN	Accuracy
Kazancli et al, 2018[22]	Spain	Clinical	59	MS	Segmentation	3D-CNN	-DSC -True Positive Rate -False Discovery Rate -Volume Difference
La Rosa et al, 2018[23]	Switzerland	Clinical	105	-Training dataset: 32 patients with EDSS scores ranged from 1 to 2 -Test dataset: 73 patients with EDSS scores ranged from 1 to 7.5	Segmentation	3D-CNN	-DSC -Lesion-wise false positive -Lesion-wise true positive -Volume difference
Roy et al, 2018[24]	USA	ISBI 2015	19	-Training dataset: 5 patients with MS -Test dataset: 14 patients with MS	Segmentation	2D-CNN	DSC
Shrwan et al, 2021[25]	India	Clinical	38	MS	Classification	2D-CNN	-Accuracy -Precision -Recall f_score
Siar et al, 2019[26]	Iran	Clinical	1111	320 MS patients 791 healthy patients	Classification	2D-CNN	-Accuracy -Sensitivity -Specificity
Valverde et al, 2018[27]	Spain	MICCAI 2008 MICCAI 2016 ISBI 2015	60	MS	Segmentation	3D-CNN	-DSC -Sensitivity -Precision
Wang et al, 2018[28]	China	eHealth Laboratory and Private data	64	38 MS patients 26 healthy patients	Identification	2D-CNN	-Accuracy -Sensitivity -Specificity
Zhang et al, 2018[29]	China	eHealth Laboratory and Private data	64	38 MS patients 26 healthy patients	Identification	3D-CNN	-Accuracy -Sensitivity -Specificity

CNN: convolutional neural network, CombiRx: Combination Therapy in Patients with Relapsing-Remitting Multiple Sclerosis, DSC: Dice Similarity Coefficient, EDSS: Expanded Disability Status Scale, ISBI: International Symposium on Biomedical Imaging, MICCAI: Medical Image Computing and Computer Assisted Intervention, MS: Multiple sclerosis.

Table 3 - Subgroup analysis.

Parameter	Number of studies	Rate of DSC [95% CI]	Heterogeneity
Country			
Australia	1	67.00 [66.80-67.20]	Chi^2 =2121.51
Switzerland	1	63.00 [62.80-63.20]	p<0.00001
USA	2	66.70 [66.56-66.83]	I^2 =100%
Spain	2	55.25 [55.11-55.39]	
Italy	2	68.17 [68.04-68.31]	
DL architecture			
2D CNN	4	64.94 [64.84-65.03]	Chi^2 =1067.22
3D CNN	4	62.63 [62.53-62.72]	p<0.00001
			I^2 =99.9%

Table 4 - Leave-one-out analysis of the rate of DSC.

Study excluded	Rate of DSC (95% CI)
Afzal et al, 2021	63.32 (57.06-69.58)
Aslani et al, 2019	62.92 (56.88-68.96)
Aslani et al, 2019	63.38 (57.10-69.66)
Coronado et al, 2020	61.89 (57.20-66.58)
Kazancli et al, 2018	64.68 (58.67-70.69)
La Rosa et al., 2018	63.89 (57.55-70.23)
Roy et al., 2018	64.84 (58.96-70.71)
Valverde et al., 2018	65.32 (60.02-70.62)

Results. ***Identification of studies.*** Literature search identified 2174 papers to be screened, of which 691 studies were duplicates and were removed. Hence, 1483 papers were screened by title and abstract and 543 were excluded for no full text article available or language other than English. Finally, 940 studies of which 940 studies were identified as potentially eligible and then were full text reviewed. Fifteen publications satisfied the eligibility requirements and were included in this study. The flowchart for the PRISMA study is shown in (**Figure 1**).

Characteristics of studies. The 15 studies were released between 2018 and 2021 and were came from nine nations: Australia (n=2), Iran (n=2), Italy (n=2), USA (n=2), China (n=2), Spain (n=2), Switzerland (n=1), India (n=1), and Germany (n=1). The number of patients ranged from 19 to 1111. Three, four, and eight studies reported the effectiveness of 2D- and 3D-CNN in the identification, classification, and segmentation of MS lesions, respectively. Study features are represented in (**Table 1**).

Table 1. Features of included studies.

For the clinical diagnosis of MS, it is crucial to identify brain lesions utilizing MRI modalities. Medical professionals have significant challenges when trying

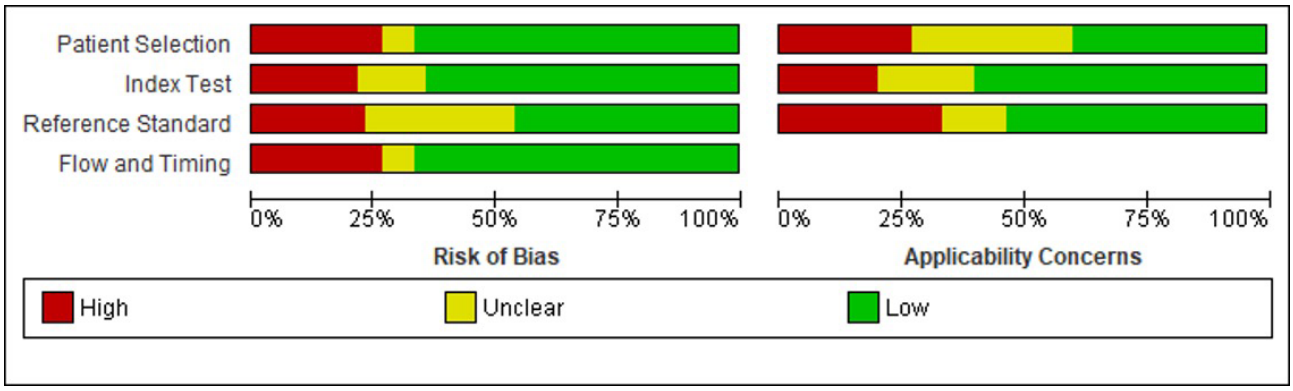

Figure 2 - Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies

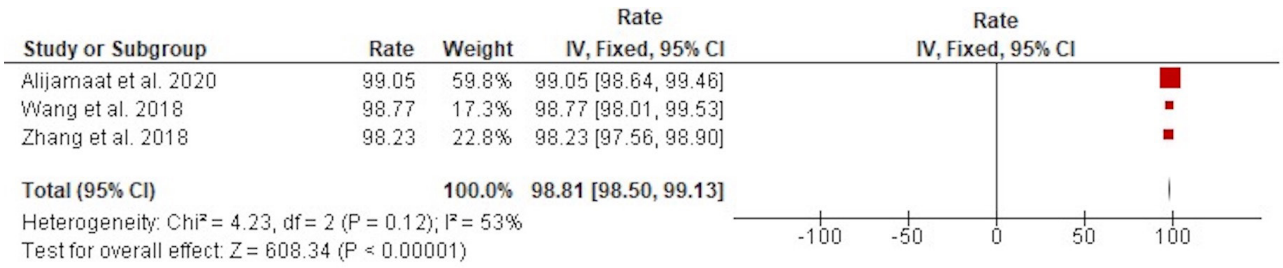

Figure 3 - Pooled accuracy rates of 2D-3D CNN in the identification of MS lesions

Table 2 - Summary of CADS developed for MS using MRI neuroimaging modalities and details of deep learning architectures.

Article	Preprocessing toolbox	Others preprocessing	Toolbox	K Fold	Details	Classifier	Loss function	Optimizer
Afzal el al., 2018[23]	-	data augmentation	Keras	-	6 convolutional layers + 6 Max Pooling	-	-	Proposed
Afzal el al., 2021[24]	FMRIB	Patch Extraction	Keras, Tensor Flow	-	2 convolutional layers + 2 Max Pooling + 1 fully connected	Multinomial LR	-	-
Alijamaat et al., 2020[25]	-	data augmentation, Histogram Stretching, discrete wavelet transform	Keras, Tensor Flow	-	15 convolutional layers + 1 Average Pooling + 1 fully connected + Dropout	Sigmoid	-	Adam
Aslani et al., 2019[26]	FMRIB	Decomposing 3D Data Into 2D Images	Keras, Tensor Flow	4	3 Parallel ResNet50s + 5 MMFF Blocks + 4 MSFU Blocks + MPR Block	Softmax	Soft Dice Loss function	Adam
Aslani et al., 2019[27]	-	Data Augmentation	Keras	-	ResNet50 + UFF Blocks	-	binary cross-entropy	Adadelta
Coronado et al., 2020[28]	-	Magnetic Resonance Imaging Automatic Processing Pipeline	-	-	5 convolutional + 4 Context Modules + 3 Up Sampling Modules + 2 Localization Modules + 2 Segmentation + 3 Strides + 3 De-Conv + 1 Upscaling	Softmax	Multiclass Weighted Dice	Adam
Eitel et al., 2019[29]	FMRIB	data augmentation	Keras, Tensor Flow	-	4 convolutional + 4 Max-Pooling + 4 Dropout	Sigmoid	-	Adam
Kazancli et al., 2018[30]	Free Surfer	Patch Extraction	Tensor Flow	-	2 convolutional + 2 Average Pooling + 2 batch normalization + 1 fully connected + 1 Dropout	Softmax	cross-entropy	Adam
La Rosa et al., 2018[31]	FMRIB	Manual Segmentation, LeMan-PV	-	-	4 convolutional + 2 Max Pooling + 4 batch normalization + 1 fully connected + 1 Dropout	Softmax	cross-entropy	Adam
Roy et al., 2018[32]	-	-	Tensor Flow, Keras	-	15 convolutional	-	-	Adam
Shrwan et al., 2021[33]	-	-	Matlab R2020a	-	3 convolutional + 3 batch normalization + 3 Max Pooling + 2 fully connected	Softmax	cross-entropy	SGDM
Siar et al., 2019[34]	-	-	-	-	25 Layers	Softmax	-	-
Valverde et al., 2018[35]	FMRIB	-	Keras, Tensor Flow	-	4 convolutional + 2 Max-Pooling + 4 batch normalization + 3 fully connected + 3 Dropout	Softmax	categorical cross-entropy	ADADELTA
Wang et al., 2018[36]	-	histogram stretching, data augmentation	-	-	11 convolutional + 11 batch normalization + 4 Pooling + 3 fully connected + 2 Dropout	Softmax	-	-
Zhang et al., 2018[37]	-	histogram stretching, data augmentation	-	-	7 convolutional +7 Pooling + 3 fully connected + 3 Dropout	Softmax	-	-

ADADELTA: adaptive learning rate method, Adam: A Method for Stochastic Optimization, CADS: Computer-aided detection software, CDMS: clinically defined multiple sclerosis, EDSS: Expanded Disability Status Scale, FMRIB: Functional Magnetic Resonance Imaging of the Brain, MMFF: multi-modal feature fusion block, , MRI: Magnetic resonance imaging, MRIAP: Magnetic Resonance Imaging Automatic Processing, MSFU: multi-scale feature upsampling block, MPR: multi-planes reconstruction, Matlab: matrix laboratory, SGDM: Stochastic Gradient Descent Momentum, UFF: upsampling fused featu

Table 2 - Summary of CADS developed for MS using MRI neuroimaging modalities and details of deep learning architectures.

Article	Clinical data about cases and controls
Afzal el al., 2018[23]	All patients included fulfilled the McDonald's criteria. Out of these 21 patients, 10 converted to CDMS after one year, whereas 11 did not convert to CDMS after one year follow up.
Afzal el al., 2021[24]	21 scans of 5 subjects are available for training purposes and already preprocessed with several steps like skull stripping, denoising, bias correction, and co-registration. These 5 subjects have 4 time points and one subject having 5 time points with a gap of approximately 1 year. These 21 scans are provided for training purposes only. For testing purposes, 61 scans are provided from 14 subjects.
Alijamaat et al., 2020[25]	MRI images of 38 MS patients whose lesions are labeled by several neurologists and approved by radiologists. To increase the number of images, MRI images of 20 healthy individuals have been prepared by the authors and added to the existing data set.
Aslani et al., 2019[26]	19 subjects divided into two sets, 5 subjects for training and 14 subjects for testing.Each subject has MRI data with a different number of time-points, normally ranging between 4 to 6.
Aslani et al., 2019[27]	37 MS patients (22 females and 15 males) with mean age 44,6±12,2 years. The patient clinical phenotypes were 24 relapsing remitting MS, 3 primary progressive MS and 10 secondary progressive MS. The mean EDSS was 3,3±2, the mean disease duration was 13.1±8,7 years and the mean lesion load was 6.2±5.7 ml.
Coronado et al., 2020[28]	-
Eitel et al., 2019[29]	76 patients with relapsing-remitting MS according to the McDonald criteria 2010 and 71 healthy controls. Patients were excluded if they were outside the age range of 18 – 69 or did not have an MRI scan. All patients were examined under supervision of a board-certified neurologist at the NeuroCure Clinical Research Center (Charité – Universitätsmedizin Berlin) between January 2011 and July 2015.
Kazancli et al., 2018[30]	-
La Rosa et al., 2018[31]	-The training dataset was composed of 32 patients, 18 female / 14 male, mean age 34±10 years, with EDSS scores ranged from 1 to 2 (mean 1,6±0,3). Mean lesion volume is 0,11±0,40 ml (range 0.001-7.03 ml). Mean lesion load per case was 6,0±7,2 ml (range 0,3-37,2 ml). -The test dataset was made up of 73 patients, 50 females and 23 males (mean age 38±10 years). EDSS scores ranged from 1 to 7.5 (mean 2,6±1,5). Mean lesion volume was 0,25±3,29 ml (range 0.002-159.827 ml). Mean lesion load per case was 14,3±27,9 ml (range 0.2-162.9 ml).
Roy et al., 2018[32]	128 patients enrolled in a natural history study of MS, 79 with relapsing-remitting, 30 with secondary progressive, and 19 with primary pro-gressive MS.
Shrwan et al., 2021[33]	-
Siar et al., 2019[34]	200 patients, including tumors and MS and healthy patients. Totally, the number of trench data for the brain tumor class was 461 images, 791 healthy patients, and 320 MS patients. The total number of data for the most 1286 images and test data was 384 images. Pictures were collected in the range of 6 to 80 years old and the average age was 43.
Valverde et al., 2018[35]	60 patients with a clinically isolated syndrome (Hospital Vall d'Hebron, Barcelona, Spain) were scanned on a 3 T Siemens with a 12-channel phased-array head coil (Trio Tim, Siemens, Germany)
Wang et al., 2018[36]	-
Zhang et al., 2018[37]	-There are 38 patients in the eHealth dataset. 676 slices associated with plaques were selected. All Brain lesions were identified and delineated by experienced MS neurologists and were confirmed by radiologists. -Age-matched and gender-matched healthy controls (HC) of the eHealth dataset were included. The exclusion criteria for all volunteers were known neurological or psychiatric diseases, brain lesions, taking psychotropic medications, and contraindications to MR imaging.

ADADELTA: adaptive learning rate method, Adam: A Method for Stochastic Optimization, CADS: Computer-aided detection software, CDMS: clinically defined multiple sclerosis, EDSS: Expanded Disability Status Scale, FMRIB: Functional Magnetic Resonance Imaging of the Brain, MMFF: multi-modal feature fusion block, , MRI: Magnetic resonance imaging, MRIAP: Magnetic Resonance Imaging Automatic Processing,, MSFU: multi-scale feature upsampling block, MPR: multi-planes reconstruction, Matlab: matrix laboratory, SGDM: Stochastic Gradient Descent Momentum, UFF: upsampling fused featu

to segment and categorize brain lesions obtained from MRI modalities and are at risk of making errors in diagnosis. Many elements, including artifacts, intensity heterogeneity, etc., have a negative impact on the MR image's quality, which frequently results in disease misdiagnosis. The low level and high level preprocessing techniques used by MRI neuroimaging modalities to diagnose MS are covered in the sections that follow. Computer aided diagnosis system (CADS) performs better when high level preprocessing techniques are used in conjunction with low level preprocessing approaches. Data augmentation (DA), patch extraction,

and other techniques are among them. **Table 2** provides a summary of the specific preprocessing data used by each article to diagnose MS utilizing DL techniques and MRI modalities. There are many toolboxes available for implementing DL models. **Table 2** lists the tools used to create DL architectures. TensorFlow and Keras are the most significant DL tools. The final component of the DL-based CADS displayed in **Table 2** is the activation function of the final layer used for classification in DL models. It can be noted that, the SoftMax function has yielded the highest classification performance.

Evaluation of the studies' quality. A high risk of bias was revealed in approximately 25% of articles regarding patient selection and flow and timing criteria. In most of the papers (75%), a certain threshold was provided in relation to the index test criteria. Moreover, in terms of reference standard, a low risk of bias was detected in less than half of the included articles. We noticed that approximately similar results were found for applicability concerns (**Figure 2**). Indeed, the highest concerns were detected in reference standard criteria (33,33%), followed by patient selection (26,67%) and index test (20%).

Types of application and outcome measures Identification. Of the 15 included studies, three studies evaluated the diagnostic effectiveness of 2D- or 3D-CNN in the identification of MS lesions using accuracy, sensitivity, and specificity (25,36,37).

Accuracy. The heterogeneity was low (Chi^2=4.23, p=0.12, I^2=53%), so a fixed effect design was used. We revealed that the overall accuracy rate was significantly high at 98.81 (95% CI: 98.50–99.13; p<0.00001) (**Figure 3**).

Sensitivity. The heterogeneity was low (Chi^2=5.17, p=0.08, I^2=61%), so we used a fixed effect model. The analysis revealed that the overall sensitivity rate was significantly high at 98.76 (95% CI: 98.42–99.10; p<0.00001) (**Figure 4**).

Specificity. The heterogeneity was low (Chi^2=1.63, p=0.44, I^2=0%), so we used a fixed effect model. The analysis revealed that the overall specificity rate was significantly high at 98.67 (95% CI: 98.22–99.12; p< 0.00001) (**Figure 5**).

Classification. Using accuracy, four of the 15 included studies examined the diagnostic efficacy of 2D- or 3D-CNN in the classification of MS lesions.[23,29,33,34]

Accuracy. The heterogeneity was important (Chi^2 = 2995.26, p<0.00001, I^2=100%), so we used a random effects model. The analysis revealed that the overall accuracy rate was significantly high at 91.38 (95% CI: 83.23–99.54; p<0.00001) (**Figure 6**).

Figure 6 pooled accuracy rates of 2D-3D CNN in the classification of MS lesions

Segmentation. Eight of the 15 included studies examined the diagnostic performance of 2D- or 3D-CNN in segmenting MS lesions using the DSC.[24,26–28,30–32,35]

Dice Similarity Coefficient. The heterogeneity was important (Chi^2=43991.74, p<0.00001, I^2 = 100%), so we used a random effects model. The analysis showed that the overall DSC was significantly high at 63.78 (95% CI: 58.29–69.27; p<0.00001) (**Figure 7**).

Subgroup and sensitivity analyses. Subgroup and sensitivity analyses were performed for DSC outcome. The small number of articles led to the exclusion of the remaining outcomes. Exploratory subgroup analysis proved that both country and DL architecture were a cause of heterogeneity for DSC outcome (p<0.00001) (**Table 3**). For the subgroup analysis of country, the highest DSC was detected in Italy (68,17%, 95%CI: 68.04, 68.3), followed by Australia (67.00%, 95%CI: 66.80, 67.20) and USA (66.70% 95%CI: 66.56, 66.83). However, the lowest DSC was revealed in Spain (55.25%, 95%CI: 55.11, 55.39). Regarding DL architecture, 2D CNN showed a higher DSC (64.94%, 95%CI: 64.84, 65.03) than 3D CNN did (62.63%, 95%CI: 62.53, 62.72).

A sensitivity analysis was carried out to determine the source of heterogeneity in the pooled rate of DSC. The finding showed that the outcomes were not significantly different between studies, suggesting that this metaanalysis is reliable. Indeed, the rate of DSC was ranged from 61.89% (95% CI: 57.20, 66.58) to 65.32% (95% CI: 60.02, 70.62) (**Table 4**).

The results of Egger's test showed that there was no publication bias for the DSC outcome (p>0.05). Similarly, the distribution of articles displayed symmetry in the funnel plot (**Figure 8**). The small number of articles led to the exclusion of the remaining out-comes. Figure 8. Funnel plot of DSC in studies investigating the segmentation of MS lesions

Discussion.

The medical sciences include many fascinating research areas, including disease prediction. The application of computer vision has led to the suggestion of numerous tools. The CNS is impacted by diseases that are ongoing, autoimmune, and demyelinating, such as MS. This harms the myelin sheath, inducing changes in the structure of the brain.[38] Consequently, it can cause disability in young people, which has a significant effect on the quality of life.[39] MS is a condition that has significant clinical implications, for which automated detecting algorithms are required to aid physicians in its early detection and faster implementation of specialist treatment. During

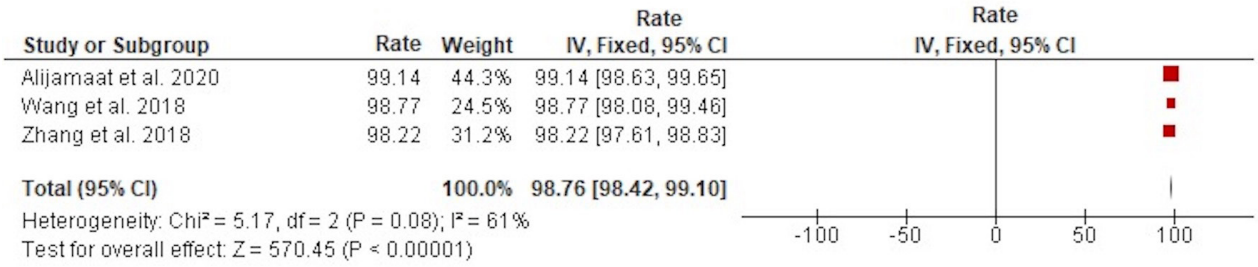

Figure 4 - Pooled sensitivity rates of 2D-3D CNN in the identification of MS lesions.

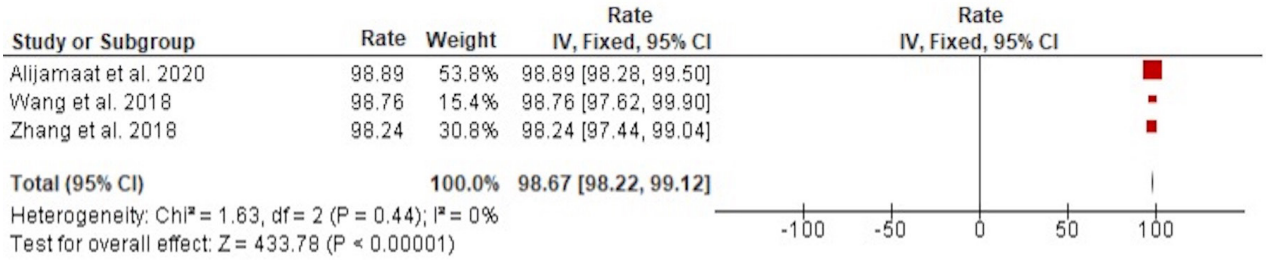

Figure 5 - Pooled specificity rates of 2D-3D CNN in the identification of MS lesions.

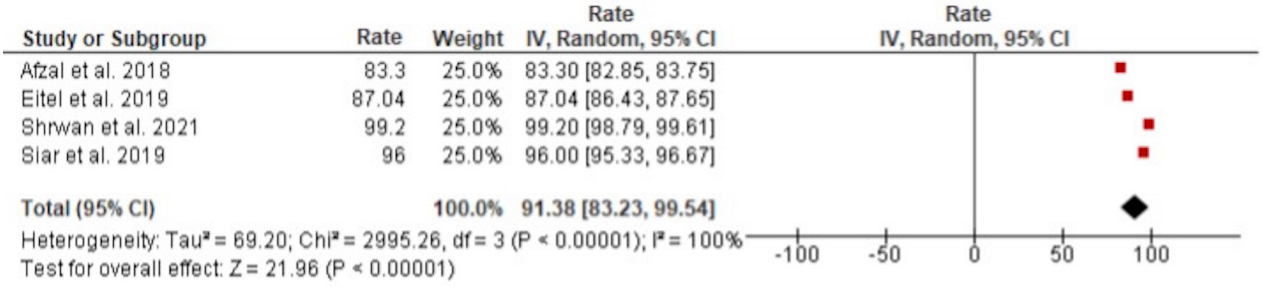

Figure 6 - Pooled accuracy rates of 2D-3D CNN in the classification of MS lesions.

the past decades, CNN, which are a ML method, have become amongst the most widely used tools in image segmentation tasks with high efficiency.[40]

The results of this analysis show that there has been an increase in interest over the past several years in using DL approaches for segmentation and classification of MS imaging investigations.

This metaanalysis evaluates the diagnostic effectiveness of 2D-3D CNN architectures using MRI data in the diagnosis of MS. The findings obtained revealed that the pooled results of 2D-3D CNN methods present a great specificity, sensitivity, and accuracy (>98%) in the identification of MS lesions. Compared to conventional techniques, they have excellent results with MS lesions.[41] For example, gray level cooccurrence matrix (GLCM) and hybrid image enhancement (HIE), which are traditional methods, presented accuracy values of 95.14% and 95.98%, respectively.[42,43] DL-based algorithms outperform conventional image processing techniques in terms of specificity, sensitivity, and accuracy requirements. As a result, 2D-3D CNN networks can effectively extract the

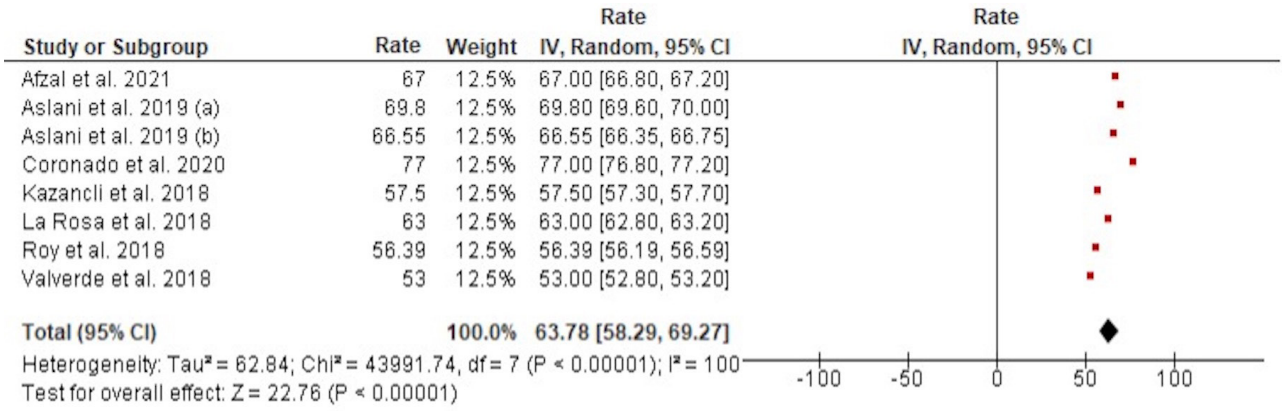

Figure 7- Pooled DSC of 2D-3D CNN in the segmentation of MS lesions.

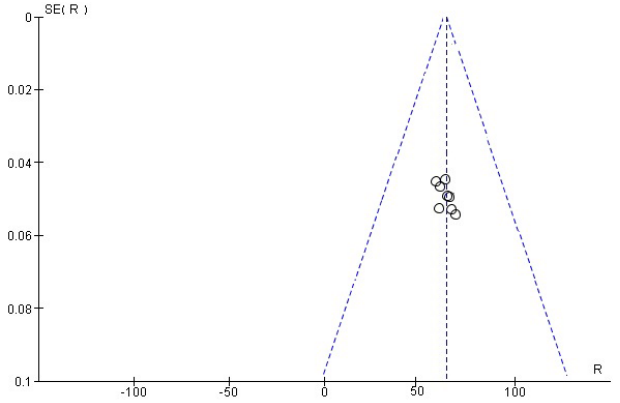

Figure 8 - Funnel plot of DSC in studies investigating the segmentation of MS lesions

features of MS lesions from MRI images. The quantity of network parameters that significantly affect network training time is another factor that can be considered in deep networks. It was revealed that 2D-3D CNN method with a smaller number of parameters provides better accuracy than other methods (traditional or DL).[25] Similarly, three techniques: stochastic pooling, dropout, and batch normalization, associated with a 14-layer CNN, were developed by Wang et al. for MS identification.[36] Compared to other cutting-edge methods, the suggested 2D-CNN method showed the best performance in terms of sensitivity, specificity, precision, and accuracy.[36]

In order to diagnose and treat MS disease as early as possible, Wang et al. constructed a 14-layer CNN together with batch normalization, dropout, and stochastic pooling algorithms.[36] During the current investigation, imaging data were collected on 26 healthy volunteers and 38 MS patients from the eHealth laboratory. Results showed that the proposed 14-layer CNN network performed better than all of the aforementioned techniques when compared to maximum pooling, average pooling, five traditional AI methods, and a deep learning method. The proposed method had sensitivity, specificity, and accuracy of 98.77%, 98.76%, and 98.77%, respectively. On the other hand, Zhang et al. proposed a new 10-layer CNN that combines dropout and parametric rectified linear unit algorithms.[37] The results outperformed four cutting-edge methods and presented sensitivity, specificity, and accuracy of 98.22%, 98.24%, and 98.23%, respectively.

Concerning classification of MS lesions, different 2D- and 3D-CNN methods were suggested to support categorization choices for clinical reviews, confirm diagnosis relevant traits, and maybe gather more knowledge about MS condition. Afzal et al[23] proposed an automated 2D-CNN algorithm using DL that can forecast whether a clinically isolated illness would develop into MS within a year of follow up. Additionally, it was computationally simple and resilient in nature. McKinley et al[44] obtained sensitivity up to 72% in MS classification on 2 separate external validation sets. Similarly, Narayana et al[45] achieved a sensitivity up to 72% on patient-basis versus sensitivity up to 78% on a slice-basis in the same study. In this context, Eitel et al. built a methodology to reveal CNN choices for classification of MS lesions based on FLAIR data and layer-wise relevance propagation.[29] Specifically, they showed that CNN models are capable of successfully separating MS patients and controls on a typically

sized neuroimaging cohort (Accuracy=87.04%). Furthermore, layer-wise relevance propagation is extremely helpful in both explaining individual network decisions as well as generally assisting in determining whether CNN models have learned significant features.

Similarly, both 2D- and 3D-CNN models showed robust segmentation of MS lesions. Rehan Afzal et al. developed a 2D patch-wise CNN that can segment MS lesions more precisely and firmly.[24] This method showed consistently higher sensitivity and precision than other traditional methods. With a precision of up to 90%, it can precisely and dependably distinguish MS lesions from images taken by different MRI scanners. As a result, doctors may automatically segment lesions without wasting time, which improves disease monitoring. In another study, Coronado et al. showed excellent segmentation by enhancing lesions using a 3D-CNN model and multispectral MRI.[28] With a testing DSC of 91%, Narayana et al.[45] offered a FLAIR based lesion segmentation, and Sander et al.[46] used a multidimensional gated recurrent unit model to achieve performance up to DSC 97%. To address the issue that MS lesions vary greatly in size and that DSC is not differentiable, making it unable to employ directly Wang et al.[47] segregate large and small lesions for gradient descent and suggest a new activation function to ease network training. Aslani et al.[27] employ 2D slices as input and a 2D encoder-decoder network to segment MS lesions in order to avoid issues like patch-wise approaches' oversight of global information and 3D segmentation's overfitting because of the class imbalance issue.

Based on the rising accessibility of larger datasets, the development of computer aided diagnosis methods for the early detection of MS disease is of utmost relevance. Such systems, which offer cloud based AI services, can improve physicians' diagnostic experiences and present improved diagnostic prospects, particularly in remote places where access to specialists is typically limited. In order to accomplish the objective of using AI algorithms in clinical practice in the actual world, all of these directions for the future of DL in MS must foster explainability and trustworthiness in addition to greater performance.

Strengths and limitations. This study investigates the diagnostic performance of CNN-based MRI in the identification, classification, and segmentation of MS lesions, taking studies from various countries into consideration. In the current investigation, we searched six different databases. The key advantages of this paper are the wide scope of datasets and the acceptable sample size. We also demonstrated the superior caliber of the included research, which had an acceptable quality grade. Our systematic review and metaanalysis have few limitations. First, there were limited papers included in this work. Moreover, a high heterogeneity was detected across papers detailing diagnostic effectiveness of 2D-3D CNN in the segmentation and classification of MS lesions. Regarding the subgroup analysis, variations in the study locations and DL architectures could be sources of heterogeneity. The results' interpretability may change as a result of substantial heterogeneity, which is expected in metaanalysis investigations.[48] Hence, the results of this metaanalysis need to be carefully considered. Despite these drawbacks, the main advantage of this work is the high methodological quality of the studies. Furthermore, the sensitivity analysis demonstrated that the calculated DSC rate was accurate and unaffected by the omission of a single study.

Conclusion. The present metaanalysis shows that 2D-3D CNN algorithms using MRI data present excellent performance in the identification, classification, and segmentation of MS lesions. However, using them directly in clinical practice is still challenging. Hence, more DL algorithms should be developed and improved for more efficacious and rapid MS diagnosis.

Acknowledgement. *The authors would like to thank Proof-Reading-Service.com (www.proof-reading-service.com) for English language editing. Further, this research holds no conflict of interest and is not funded through any source.*

References

1. Walton C, King R, Rechtman L, Kaye W, Leray E, Marrie RA, et al. Rising prevalence of multiple sclerosis worldwide: Insights from the Atlas of MS, third edition. ***Mult Scler déc*** 2020; 26: 1816-1821.
2. Howard J, Trevick S, Younger DS. Epidemiology of Multiple Sclerosis. ***Neurologic Clinics*** 2016; 34: 919-939.
3. Calabresi PA. Diagnosis and management of multiple sclerosis. ***Am Fam Physician*** 2004; 70: 1935-1944.
4. Eshaghi A, Young AL, Wijeratne PA, Prados F, Arnold DL, Narayanan S, et al. Identifying multiple sclerosis subtypes using unsupervised machine learning and MRI data. ***Nat Commun*** 2021; 12: 2078.
5. Hartmann M, Fenton N, Dobson R. Current review and next steps for artificial intelligence in multiple sclerosis risk research. ***Comput Biol Med*** 2021; 132: 104337.
6. Sawcer S, Franklin RJM, Ban M. Multiple sclerosis genetics. ***Lancet Neurol*** 2014; 13: 700-709.

7. Landfeldt E, Castelo-Branco A, Svedbom A, Löfroth E, Kavaliunas A, Hillert J. The long-term impact of early treatment of multiple sclerosis on the risk of disability pension. ***J Neurol*** 2018; 265: 701-707.
8. Buyukturkoglu K, Zeng D, Bharadwaj S, Tozlu C, Mormina E, Igwe KC, et al. Classifying multiple sclerosis patients on the basis of SDMT performance using machine learning. ***Mult Scler*** 2021; 27: 107-116.
9. Filippi M, Rocca MA, Ciccarelli O, De Stefano N, Evangelou N, Kappos L, et al. MRI criteria for the diagnosis of multiple sclerosis: MAGNIMS consensus guidelines. ***Lancet Neurol*** 2016; 15: 292-303.
10. Richards TL. Proton MR spectroscopy in multiple sclerosis: value in establishing diagnosis, monitoring progression, and evaluating therapy. ***AJR Am J Roentgenol*** 1991; 157: 1073-1078.
11. Filippi M, Rocca MA. Magnetization Transfer Magnetic Resonance Imaging in the Assessment of Neurological Diseases. ***J Neuroimaging*** 2004; 14: 303-313.
12. Rovaris M, Gass A, Bammer R, Hickman SJ, Ciccarelli O, Miller DH, et al. Diffusion MRI in multiple sclerosis. ***Neurology*** 2005; 65: 1526-1532.
13. Lapointe E, Li DKB, Traboulsee AL, Rauscher A. What Have We Learned from Perfusion MRI in Multiple Sclerosis? ***AJNR Am J Neuroradiol*** 2018; 39: 994-1000.
14. Shoeibi A, Khodatars M, Jafari M, Moridian P, Rezaei M, Alizadehsani R, et al. Applications of deep learning techniques for automated multiple sclerosis detection using magnetic resonance imaging: A review. ***Comput Biol Med*** 2021; 136: 104697.
15. Górriz JM, Ramírez J, Ortíz A, Martínez-Murcia FJ, Segovia F, Suckling J, et al. Artificial intelligence within the interplay between natural and artificial computation: Advances in data science, trends and applications. ***Neurocomputing*** 2020; 410: 237-270.
16. Martinez-Murcia FJ, Górriz JM, Ramírez J, Ortiz A. Convolutional Neural Networks for Neuroimaging in Parkinson's Disease: Is Preprocessing Needed? ***Int J Neural Syst*** 2018; 28: 1850035.
17. Sadeghi D, Shoeibi A, Ghassemi N, Moridian P, Khadem A, Alizadehsani R, et al. An Overview on Artificial Intelligence Techniques for Diagnosis of Schizophrenia Based on Magnetic Resonance Imaging Modalities: Methods, Challenges, and Future Works. ***Comput Biol Med*** 2022: 146: 105554.
18. Jiménez-Mesa C, Ramírez J, Suckling J, Vöglein J, Levin J, Górriz JM, et al. Deep Learning in current Neuroimaging: a multivariate approach with power and type I error control but arguable generalization ability. ***ArXiv*** 2103: 16685.
19. McKinley R, Wepfer R, Aschwanden F, Grunder L, Muri R, Rummel C, et al. Simultaneous lesion and brain segmentation in multiple sclerosis using deep neural networks. ***Sci Rep*** 2021; 11: 1087.
20. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. ***BMJ*** 2021 29; 372.
21. Whiting PF. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. ***Ann Intern Med*** 2011; 155: 529-536.
22. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. ***Res Synth Method*** 2010; 1: 97-111.
23. Afzal H, Luo S, Ramadan S, Lechner-Scott J, Li J. Automatic Prediction of the Conversion of Clinically Isolated Syndrome to Multiple Sclerosis Using Deep Learning. In: International conference on video and image processing. 2018: 231-235. From: https://doi.org/10.1145/3301506.3301526
24. Afzal HM, Luo S, Ramadan S, Lechner-Scott J, Amin MR, Li J, et al. Automatic and Robust Segmentation of Multiple Sclerosis Lesions with Convolutional Neural Networks. ***Computers, Materials & Continua*** 2021; 66: 1.
25. Alijamaat A, NikravanShalmani A, Bayat P. Multiple sclerosis identification in brain MRI images using wavelet convolutional neural networks. ***Int J Imaging Syst Technol*** 2021; 31: 778-785.
26. Aslani S, Dayan M, Storelli L, Filippi M, Murino V, Rocca MA, et al. Multi-branch convolutional neural network for multiple sclerosis lesion segmentation. ***Neuroimage*** 2019; 196: 1-15.
27. Aslani S, Dayan M, Murino V, Sona D. Deep 2D Encoder-Decoder Convolutional Neural Network for Multiple Sclerosis Lesion Segmentation in Brain MRI. In: Crimi A, Bakas S, Kuijf H, Keyvan F, Reyes M, van Walsum T, éditeurs. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries (Internet). Cham: Springer International Publishing; 2019 (cité 6 déc 2021). p. 132-41. (Lecture Notes in Computer Science; vol. 11383). Disponible sur: http://link.springer.com/10.1007/978-3-030-11723-8_13
28. Coronado I, Gabr RE, Narayana PA. Deep learning segmentation of gadolinium-enhancing lesions in multiple sclerosis. ***Mult Scler*** 2021; 27: 519-527.
29. Eitel F, Soehler E, Bellmann-Strobl J, Brandt AU, Ruprecht K, Giess RM, et al. Uncovering convolutional neural network decisions for diagnosing multiple sclerosis on conventional MRI using layer-wise relevance propagation. ***Neuroimage Clin*** 2019; 24: 102003.
30. Kazancli E, Prchkovska V, Rodrigues P, Villoslada P, Igual L. Multiple Sclerosis Lesion Segmentation using Improved Convolutional Neural Networks: In: Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (Internet). Funchal, Madeira, Portugal: SCITEPRESS - Science and Technology Publications; 2018 (cité 6 déc 2021). p. 260-9. Disponible sur: http://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0006540902600269
31. La Rosa F, Fartaria MJ, Kober T, Richiardi J, Granziera C, Thiran JP, et al. Shallow vs Deep Learning Architectures for White Matter Lesion Segmentation in the Early Stages of Multiple Sclerosis. In: Crimi A, Bakas S, Kuijf H, Keyvan F, Reyes M, van Walsum T, éditeurs. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries (Internet). Cham: Springer International Publishing; 2019 (cité 6 déc 2021). p. 142-51. (Lecture Notes in Computer Science; vol. 11383). Disponible sur: http://link.springer.com/10.1007/978-3-030-11723-8_14
32. Roy S, Butman JA, Reich DS, Calabresi PA, Pham DL. Multiple sclerosis lesion segmentation from brain MRI via fully convolutional neural networks. ***ArXiv*** 2018; 1803.09172.
33. Shrwan R, Gupta A. Classification of Pituitary Tumor and Multiple Sclerosis Brain Lesions through Convolutional Neural Networks. I***OP Conf Ser: Mater Sci Eng*** 2021; 1049: 012014.
34. Siar H, Teshnehlab M. Diagnosing and classification tumors and MS simultaneous of magnetic resonance images using convolution neural network. In2019 7th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS). IEEE: Piscataway (NJ) USA; 2019. p. 1-4.

35. Valverde S, Salem M, Cabezas M, Pareto D, Vilanova JC, Ramió-Torrentà L, et al. One-shot domain adaptation in multiple sclerosis lesion segmentation using convolutional neural networks. ***NeuroImage: Clinical*** 2019; 21: 101638.
36. Wang SH, Tang C, Sun J, Yang J, Huang C, Phillips P, et al. Multiple Sclerosis Identification by 14-Layer Convolutional Neural Network With Batch Normalization, Dropout, and Stochastic Pooling. ***Front Neurosci*** 2018; 12: 818.
37. Zhang YD, Pan C, Sun J, Tang C. Multiple sclerosis identification by convolutional neural network with dropout and parametric ReLU. ***Journal of Computational Science*** 2018; 28: 1-10.
38. Zhao Y, Guo S, Luo M, Shi X, Bilello M, Zhang S, et al. A level set method for multiple sclerosis lesion segmentation. ***Magn Reson Imaging*** 2018; 49: 94-100.
39. Lladó X, Oliver A, Cabezas M, Freixenet J, Vilanova JC, Quiles A, et al. Segmentation of multiple sclerosis lesions in brain MRI: A review of automated approaches. ***Information Sciences*** 2012; 186: 164-185.
40. Anwar SM, Majid M, Qayyum A, Awais M, Alnowami M, Khan MK. Medical Image Analysis using Convolutional Neural Networks: A Review. ***J Med Syst*** 2018; 42: 226.
41. Danelakis A, Theoharis T, Verganelakis DA. Survey of automated multiple sclerosis lesion segmentation techniques on magnetic resonance imaging. ***Comput Med Imaging Graph*** 2018; 70: 83-100.
42. Ghribi O, Sellami L, Ben Slima M, Ben Hamida A, Mhiri C, Mahfoudh KB. An Advanced MRI Multi-Modalities Segmentation Methodology Dedicated to Multiple Sclerosis Lesions Exploration and Differentiation. ***IEEE Trans Nanobioscience*** 2017; 16: 656-665.
43. Ullah Z, Farooq MU, Lee SH, An D. A hybrid image enhancement based brain MRI images classification technique. ***Med Hypotheses*** 2020; 143: 109922.
44. McKinley R, Wepfer R, Grunder L, Aschwanden F, Fischer T, Friedli C, et al. Automatic detection of lesion load change in Multiple Sclerosis using convolutional neural networks with segmentation confidence. ***NeuroImage: Clinical*** 2020; 25: 102104.
45. Narayana PA, Coronado I, Sujit SJ, Wolinsky JS, Lublin FD, Gabr RE. Deep Learning for Predicting Enhancing Lesions in Multiple Sclerosis from Noncontrast MRI. ***Radiology*** 2020; 294: 398-404.
46. Sander L, Pezold S, Andermatt S, Amann M, Meier D, Wendebourg MJ, et al. Accurate, rapid and reliable, fully automated MRI brainstem segmentation for application in multiple sclerosis and neurodegenerative diseases. ***Hum Brain Mapp*** 2019; 40: 4091-4104.
47. Wang Z, Smith CD, Liu J. Ensemble of multi-sized FCNs to improve white matter lesion segmentation. ***Springer International Publishing*** 2018; 11046: 223-232.
48. Imrey PB. Limitations of Meta-analyses of Studies With High Heterogeneity. ***JAMA Netw Open*** 2020; 3: e1919325.