Received: 19 September 2024 | Revised: 26 November 2024 | Accepted: 2 January 2025 | Published online: 12 February 2025

RESEARCH ARTICLE

Journal of Data Science and Intelligent Systems
2025, Vol. 00(00) 1–11
DOI: 10.47852/bonviewJDSIS52024391

BON VIEW PUBLISHING

Digital Twin-Assisted Deep Learning with Model Fusion for Detecting Multiple Sclerosis in MRI Modalities

Ramya Palaniappan[1,]* and Siva Rathinavelayutham[1]

[1]*Department of Computational Intelligence, SRM Institute of Science and Technology, India*

Abstract: The advancement of technologies such as IoT, Big Data, Data Science, Augmented Reality/Virtual Reality, and cloud computing has transformed the manufacturing sector through the Digital Twin (DT), serving as a potent instrument for simulating concepts into practice. Recently, artificial intelligence has been merged with digital technology to conduct research in the healthcare sector, facilitating judgments on the planning of patients' clinical courses and the allocation of available medical resources. DTs can be utilized in the medical industry to facilitate clinical decision-making, providing prognoses and personalized treatment for patients. Multiple sclerosis (MS) is an autoimmune neurological disorder that impacts the central nervous system, potentially resulting in neurological impairment and mortality if left untreated. [X1]. In this research study, magnetic resonance imaging (MRI) samples collected from the E-Health lab and IXI databases are used to construct the DT empowered with artificial intelligence for the diagnosis of MS in a robust manner. We propose a hybrid CNN-RNN model for detecting MS in two phases. In the first phase, the deep features of the MRI modalities are extracted by two transfer learning models, m-InceptionV3 and m-DenseNet121. In the second phase, the classification of extracted features to either healthy or MS is performed by the Long Short-Term Memory RNN model. Deep learning metrics like precision, recall, $F1$ score, and accuracy are used to validate the performance of the proposed model. The proposed model outperformed the other state-of-the-art models achieving a good performance of 99.67% in validation accuracy. This healthcare DT pipeline may assist in clinical decision-making for MS detection and planning post-MS rehabilitation.

Keywords: multiple sclerosis, digital twin, artificial intelligence, precision medicine, precision treatment, machine learning, rehabilitation

1. Introduction

Multiple sclerosis (MS) is a severe demyelinating disorder that impacts the central nervous system, encompassing the brain, spinal cord, and optic nerves. The nerve fibers and the cells responsible for producing the myelin sheath in the brain are damaged in individuals with MS. If ignored, this injury will affect the spinal cord, potentially resulting in gait abnormalities and physical handicap as the condition advances [1]. The impairment of cells and nerve fibers manifests as lesions in various regions of the brain's gray and white matter, thus the designation "MS".

The National Multiple Sclerosis Society (2022) reports that more than 2.6 million individuals globally are afflicted with MS. Globally, around 280 individuals are diagnosed with MS daily, with a new diagnosis occurring every five minutes, according to the Atlas of MS study, 2022. The Multiple Sclerosis International Federation indicates that the nations with the highest prevalence of MS are predominantly located in northern Europe, specifically Norway, Sweden, Denmark, and Iceland. Countries with a significant prevalence of MS include Canada, the United States, Australia, and New Zealand.

*Corresponding author: Ramya Palaniappan, Department of Computational Intelligence, SRM Institute of Science and Technology, India. Email: rp3516@srmist.edu.in

MS infection comprises four stages: (i) Clinically isolated syndrome (CIS) – This is the initial stage of MS, characterized by minimal damage to nerve fibers and the myelin sheath, sometimes undetectable in a magnetic resonance imaging (MRI) scan [2]; (ii) Relapsing-Remitting MS – In this phase, symptoms exacerbate and subsequently improve, often adhering to a consistent pattern that will eventually progress to Secondary Progressive MS (SP-MS); (iii) Primary Progressive MS – This variant of MS is characterized by a continuous progression of the disease without periods of remission. This indicates heightened neurological impairment and physical incapacity; and (iv) SP-MS – It is the most aggressive form of MS, and it often happens within 10 to 25 years of the first diagnosis. If left untreated, it may progress to permanent physical handicap or possibly result in the individual's death.

Currently, there are no clinical manifestations or laboratory evaluations that may precisely identify MS. Consequently, multiple approaches were implemented for diagnosing MS, including the examination of the patient's medical history, analysis of sMRI and fMRI modalities, cerebrospinal fluid analysis, and blood testing. Currently, MRI modalities are regarded as the most effective non-invasive method for detecting MS. In a MRI scan, MS activity manifests as either hyperintense or hypointense lesions. Oval or frame-shaped lesions are prevalent among patients with MS. Both the white and gray matter of the brain can manifest MS lesions. Gadolinium is a chemical contrast agent utilized by medical professionals to augment the contrast of MRI scan images.

Modern healthcare leans towards a preventive, multidisciplinary approach to deliver personalized and precise treatment to people. The advent of artificial intelligence (AI) facilitated the amalgamation of complex insights on human physiology, behavior, and medicine. Digital twin (DT) models for healthcare have been suggested in various medical domains, including the development of DTs for the liver [1] and the heart [3, 4]. Considering the dominance of MRI procedures and the scarcity of radiologists and neurologists, the need to build an AI-based tool for accurate and early diagnosis of MS is inevitable. The developed Healthcare Digital Twin (HDT) model will aid clinicians in detecting MS.

The following are the major contributions of the research work:

1) MRI modalities from large EHealth Lab and IXI datasets are collected and pre-processed for use in this study.
2) A DT framework enabled with AI for detecting MS is proposed.
3) Deep features from MRI samples are extracted using two transfer learning models with customized layers. The extracted features are stacked up and given as input to the Recurrent Neural Network (RNN) model to make the final prediction.
4) Measured the performance of the trained hybrid CNN-RNN model with the training dataset and achieved ideal accuracy, $F1$ score, precision, recall, and ROC.

The remaining sections of the paper are organized as mentioned. Section 2 analyzes the associated research works for MS detection using deep learning and the DT concept. Section 3 highlights the enabling technologies for building DT. Section 4 describes the proposed DT framework based on InceptionV3 and DenseNet121 CNN models and the Long Short-Term Memory (LSTM) model. Section 5 presents the experimental setup, datasets, performance metrics, discusses the results, and compares them with the state-of-the-art methods. Finally, Section 6 concludes the paper and plans the future research opportunities.

2. Related Work

This research work introduces CloudDTH, a novel and extensible cloud healthcare system framework that integrates DT technology for diagnosing individual health, particularly focusing on elderly healthcare. The framework addresses the challenges of managing the health of elderly patients and integrating the medical physical world with the virtual world for smart healthcare solutions. CloudDTH utilizes technologies like big data, cloud computing, and IoT, emphasizing precision medicine, and efficient service delivery. The proposed DT healthcare (DTH) developed by Liu et al. [5] concept and model forms the foundation of CloudDTH, enabling interaction between physical and virtual spaces. The paper also explores key enabling technologies, outlines a reference framework, and demonstrates the feasibility through application scenarios and a real-time supervision case study. Laamarti et al. [6] introduce an ISO/IEEE 11073 standardized DT pipeline for well-being of senior citizens, expanding the application of DT technology beyond industry into smart healthcare services for smart cities. The framework involves a cyclic process, which includes gathering healthcare data from personal health analyzing IoT devices, analyzing them, and providing feedback to the user. Notably, the framework accommodates both X73-compliant and noncompliant healthcare devices through the integration of an X73 wrapper module. Additionally, a configurable X73 mobile application is proposed, designed to interface with any X73-compliant device. The paper describes the design and implementation of this framework, along with a proof-of-concept experiment, demonstrating effective results and the potential for gaining valuable insights into individual health while providing feedback to users and caregivers.

In this research study, Hussain et al. [7] explore the use of EEG monitoring as a diagnostic tool for neurological impairment in stroke patients, aiming to create a proof-of-concept healthcare "DT". The research, involving 48 stroke patients and 75 healthy individuals, utilized portable EEG devices to capture data from various cortical electrodes. Key characteristics identified for classification included the revised brain-symmetry index, theta, and delta activities in both motor and cognitive states. Applying support vector machine in a machine-learning approach, the study achieved a 76% accuracy in classifying EEG features, effectively distinguishing between stroke patients and the control group. The proposed healthcare DT framework, based on EEG data and machine learning, holds promise for aiding clinical decisions for stroke patients. Overall, the findings suggest a potential application of DT technology in enhancing diagnostics and personalized care for individuals with neurological impairments, particularly in the context of stroke management. This study addresses the urgent need for early detection of chronic kidney disease using deep learning utilizing CT scan images of the abdomen and urogram. The dataset comprises 12,446 images, with varying distributions for cysts, normal cases, kidney stones, and tumors. Sasikaladevi and Revathi [8] develop a deep learning model where complex features are extracted from these images, forming hypergraphs for representation. The hypergraphs are utilized in a hypergraph convolutional neural network, showcasing superior performance with a validation accuracy of 99.71%. Validation metrics include precision, recall, accuracy, and the $F1$ score, demonstrating the model's robustness. The proposed digital twin model proves to be a promising tool for nephrologists in the early diagnosis and prognosis of kidney diseases. Ojo et al. [9] introduce FAD, a hybrid deep neural network and artificial neural network model, to detect MS using gene expression data from the GEO GSE17048 dataset. Pre-processing included encoding, scaling, and feature selection, yielding a 96.55% accuracy and 96.71% $F1$ score, outperforming prior methods. Zhang et al. [10] employed enhanced convolutional neural network for detecting MS from brain slices. On training the proposed model with 676 MRI modalities of MS patients and 681 MRI modalities of healthy control subjects, the model incorporates data augmentation, parametric rectified linear unit (PReLU), and dropout approaches for improving the performance. The final 10-layer deep convolutional neural network achieves exceptional performance, with a sensitivity of 98.21%, a specificity of 98.25%, and an accuracy of 98.21%, outperforming four related approaches. Notably, the inclusion of dropout and PReLU techniques contributes to significant accuracy improvements compared to conventional methods.

Özkaraca et al. [11] proposed a novel modular deep learning model developed to enhance the classification of MRI images for chronic nerve ailments like brain tumors, strokes, dementia, and MS. Leveraging transfer learning techniques from established methods like DenseNet, VGG16, and basic CNN architectures, the model aimed to improve classification performance while addressing inherent limitations. Evaluation using both traditional train-test splits (80:20) and cross-validation (10 folds) demonstrated superior performance compared to existing transfer learning methods, albeit with increased processing time. Jain et al. [12] proposed an ensemble-based classification framework to classify healthy individuals from MS patients using MRI samples from the E-Health Lab data repository. Feature Extraction is performed using the Gray Level Co-occurrence matrix and three boosting techniques are performed on the Decision Tree Classifier achieving an accuracy of 94.91% when compared to previous related works. Öztürk and Özkaya [13] built a hybrid model for classifying gastrointestinal infections by combining CNN with RNN. The model performance is equated with various pre-trained models like AlexNet, Google Net, and ResNet architectures to assess its classification performance.

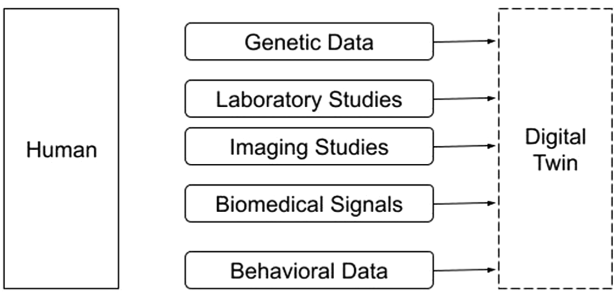
Ekmekyapar and Taşçı [14] proposed the MobileNetV2 model combined with IMrMr feature selection and KNN Nearest Neighbors classification method to detect MS. The method is conducted in two different datasets, comprising 1373 and 1385 MRI samples, respectively. The proposed model attained an accuracy efficacy of 98.76% on applying exemplar-based learning to the MobileNetV2 model. Ambayiram and Ganesan [15] developed a hybrid deep convolution network (HD-CNN) for MS classification. The model employed the Chaotic Leader-Selective Particle Swarm Optimization method for extracting white regions in brain MRI. Later on, Refined Slime Mold selects salient features needed for better segmentation. Finally, the maximum bidirectional gradient classifier categorizes MS lesions by detecting white matter spots in the target region.

3. Enabling Technologies for Virtualizing Patient's Medical Condition

3.1. Digital twin

A DT is a dynamic, changing, and intelligent digital representation of a physical object or process that is frequently likened to a "Product Avatar". While DT has gained recognition in manufacturing for enhancing failure management, boosting productivity, and streamlining processes to cut costs, its application in healthcare is still in the early stages and requires substantial research [16]. Building a DT for a patient's health condition involves two pivotal components as depicted in Figure 1 [17]. First, as requested by healthcare professionals, the physical realm contains a variety of data types, such as clinical records, laboratory test results, and patient-reported outcome measures, which are carefully gathered during the many stages of MS therapy. Subsequently, these reports are transmitted to the digital realm, where the patient's DT is crafted, leveraging AI to analyze and propose clinical treatments tailored to MS patients. Finally, these recommendations are relayed back to the physical world to guide healthcare interventions and strategies for optimal patient care.

Figure 1
Digital twinning of patient health condition for planning treatment



3.2. Role of artificial intelligence

AI, IoT, AR/VR, wireless technologies, and other fields have all contributed to recent developments in DT technology. Future DT research and development will be greatly influenced by these areas. A model that combines DT and AI was presented by Derraz et al. [18] to suggest individualized medication regimens for everyone. Similarly, Angulo et al. [19] developed a DT model to monitor lung cancer progression in patients, offering tailored recommendations. Integrating DT with AI creates opportunities to generate personalized recommendations based on the patient's actual health status, enabling doctors to make more accurate, individualized, and efficient clinical decisions for improved treatment quality.

4. Proposed Methodology

The proposed HDT for detecting MS infection is shown in Figure 2. The workflow involves several components which include dataset collection, data preparation and pre-processing, splitting the data for training, validation, and testing, building CNN models m-InceptionV3, m-DenseNet101 on applying transfer learning for feature extraction, stacking up of extracted features to train RNN classifier, LSTM, and performance measure. The input data from the IXI database and E-Health lab repository undergo various image preparation and pre-processing procedures to make it suitable for training the CNN models. Subsequently, the pre-processed image data is fed as input to two CNN models m-Inception-V3 and m-DenseNet121 which serve as the base models for applying the transfer learning approach. To adapt the pre-trained models to the specific task of MS detection, additional layers are added to the top of the models upon removing the top layer. These layers will extract the features relevant to MS detection. The features from the two transfer learning models are then stacked up and fed as input to the LSTM model to make the final classification. By combining the features of multiple models to train the meta-classifier, this approach aims to increase the accuracy and robustness of MS detection. As new patient data emerge, the DTs will adjust and refine to align with the specific conditions of the patients, thereby enhancing the accuracy and relevance of the model. This DT model will be available for clinicians to enhance their clinical decisions and follow up the treatment outcomes.

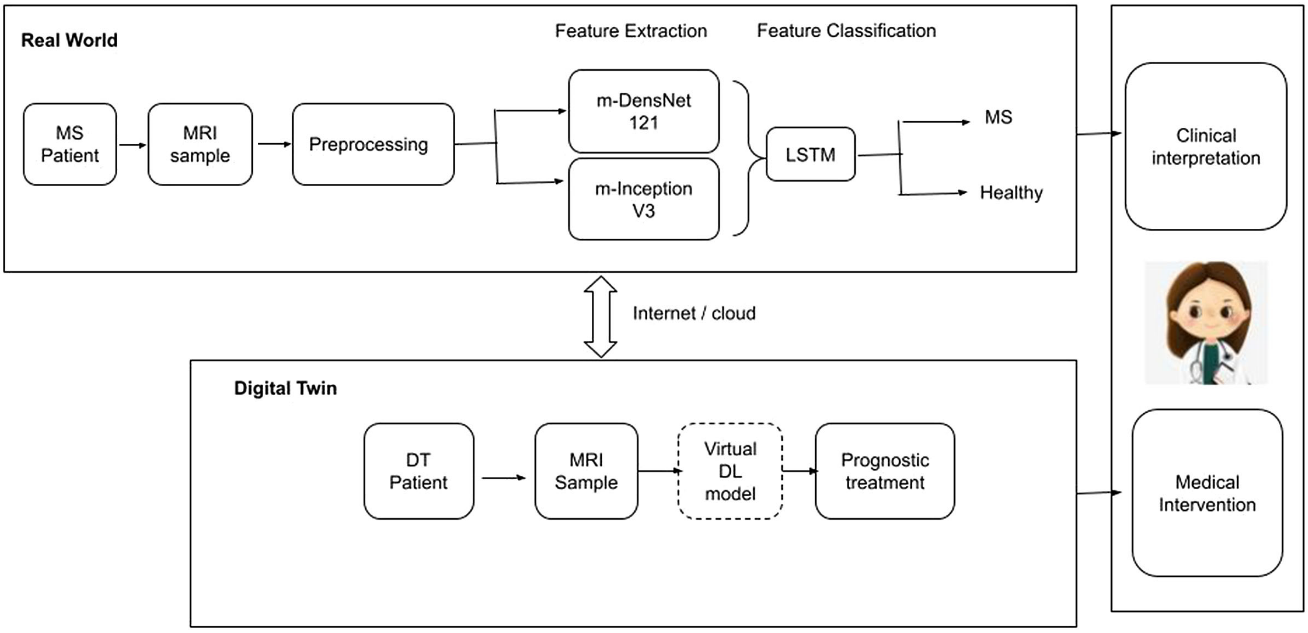
4.1. Data pre-processing

Data pre-processing is a crucial step to prepare the input data more suitable for training deep learning models. In this research work, the MRI modalities are collected from two unique datasets EHealth lab [21] and IXI dataset [21] which vary in format, dimensions, and resolution. Therefore, several pre-processing procedures are applied which include 3D-2D conversion, resizing, normalization, and data augmentation.

4.1.1. 3D to 2D conversion

In the IXI dataset, each MRI NIfTI volume consists of 256 slices. To facilitate analysis, these volumes were initially converted to PNG format using a Python code leveraging the nifti2png package. Subsequently, for each MRI volume, four specific slices were selected based on the clarity of the substantia nigra. The conversion process to PNG format enables easier visualization and manipulation of the MRI data, which is essential for subsequent analysis.

Figure 2
Envisioned workflow of Healthcare Digital Twin (HDT) for detecting MS infection



4.1.2. Image resizing and normalization

To avoid dimension mismatch and ensure consistent input, all the MRI samples are resized to a fixed dimension of 224 × 224. The pixel values of all the collected MRI modalities are adjusted to be within the standard range of 0–1 using the min-max scaling procedure. The normalized images improved the training stability and convergence of the neural network built.

4.2. Transfer learning

Transfer learning is a machine-learning procedure where knowledge gained by a model in a specific task is applied to solve another, associated task. It typically involves utilizing a pre-trained model, such as VGG16, ResNet, or Inception, which has already learned to extract meaningful features from input data. To adapt the pre-trained model to a new but related problem, additional layers such as the convolution layer, dense layer, dropout layer, and pooling layer are added to the existing model, or existing layers are fine-tuned to extract complex features from the new input data.

4.3. Modified Inception V3 network architecture (m-InceptionV3)

Inception V3, the third iteration of Inception architecture, was pre-trained on the large image dataset ImageNet [20], making it a widespread option for transfer learning-related applications. The InceptionV3 architecture is distinguished by its use of multiple parallel convolutional layers of varying sizes within a single layer. This design enables the network to capture features at multiple scales simultaneously. The input dimension for the model is adjusted to 224 × 224 × 3 to match the pre-processed MRI images consistently. The top layer of the pre-trained network is removed, and additional layers are added to customize the learning for MS detection as depicted in Figure 3. The extracted features are then fed into the GlobalAveargePooling layer to reduce the dimensionality of features. Later on, the dimensionality-reduced features pass through a fully connected layer comprising of 1024 neurons which learn complex patterns in the data by applying a series of linear and non-linear transformations. Subsequently, a dropout of 0.5 is connected to randomly drop neurons while learning to avoid model overfitting. The output is then connected to another dense layer comprising of 512 neurons.

4.4. Modified denseNet121 network architecture (m-DenseNet121)

Dense Net, including the DenseNet-121 architecture, was introduced by Huang et al. [21] in 2017 which specializes in the dense connectivity pattern in which each layer is linked to every single other layer in a feed-forward manner. This connectivity pattern fosters feature reuse and facilitates gradient flow throughout the network, addressing the vanishing gradient problem commonly encountered in deep networks. The top layer of the pre-trained model is removed, and additional layers are added to enhance the learning for MS detection as depicted in Figure 4.

4.5. LSTM model

The LSTM Model is a kind of RNN, widely used in deep learning. A LSTM unit, with its input, forget, and output gates and memory cell, resembles a layer of neurons in a traditional feed-forward neural network, albeit with specialized mechanisms for handling sequential data. Each LSTM unit has three gates. First, the forget gate for deciding which information should be remembered and which irrelevant information can be dropped. Second, the input gate is responsible for learning new features from the input, and the third one is the output gate that computes the final output after combining the outputs from the other two gates. The extracted features from the CNN module, comprising m-InceptionV3 and m-DenseNet121 models, are flattened and reshaped before being fed into the RNN module. This RNN module consists of 150 LSTM cells utilizing a rectified linear

Figure 3
m-InceptionV3 architecture for feature extraction

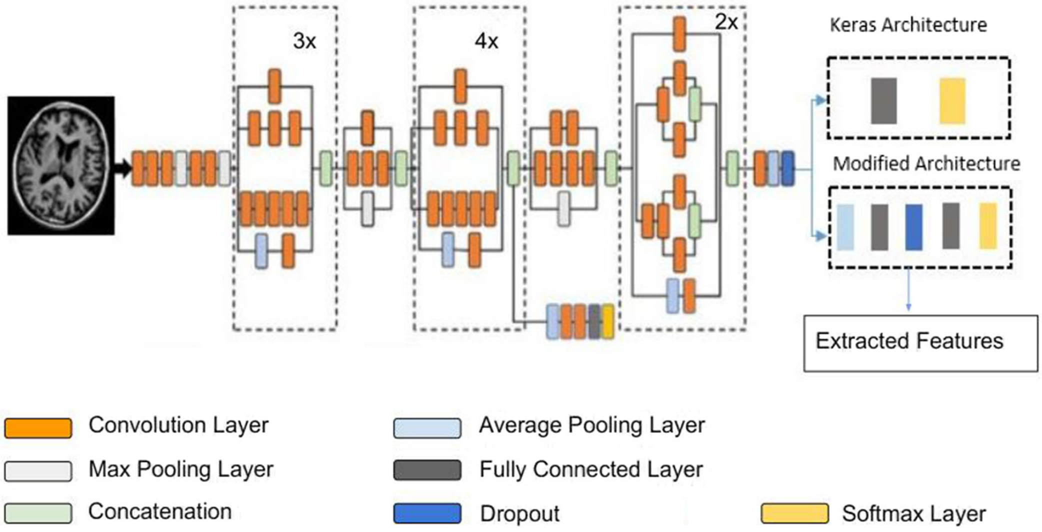
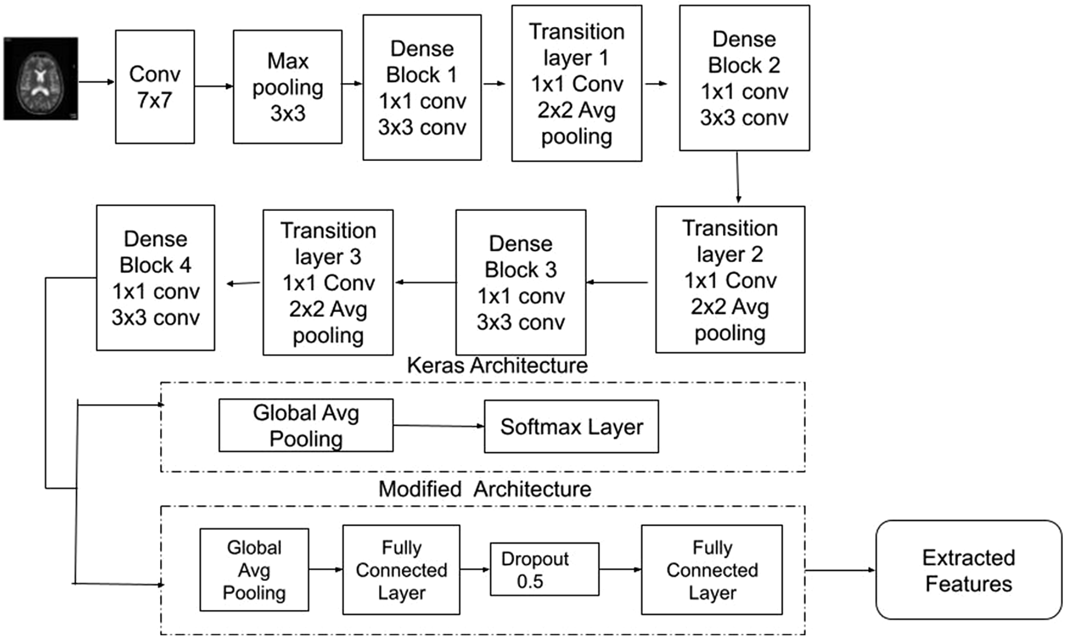


Figure 4
m-DenseNet121 architecture for feature extraction



activation function (ReLU) and Batch Normalization. Subsequently, a Flatten layer is added, followed by two blocks of Dense layers with ReLU activation and Batch Normalization layers to improve the training process. To prevent overfitting, a Dropout layer is inserted between the two Dense layers. Finally, the output layer is a Dense layer with 2 units and a Softmax activation function. The architecture of the proposed LSTM model with custom layers is depicted in Figure 5.

4.6. Stacked ensemble learning

The concept of ensemble learning is that training data is analyzed by multiple deep learning models to extract features. These features are then stacked up and utilized as meta-features to train a final ensemble classifier. We use a hybrid CNN-RNN model in our study which has two stages. In the first stage, two CNN models extract features from the training and validation data. These features are then stacked up to create complex feature representations. In the second stage, resulting features which capture essential information from the original high-dimensional feature space are then utilized as meta-features for training a final RNN classifier. Deep learning models (also referred to as first-stage CNN models) and a meta-learner (or, final-stage RNN classifier) that learns from the CNN model predictions make up the hybrid model. The algorithm of the proposed hybrid model is presented in appendix section.

Figure 5
Proposed LSTM architecture for feature classification

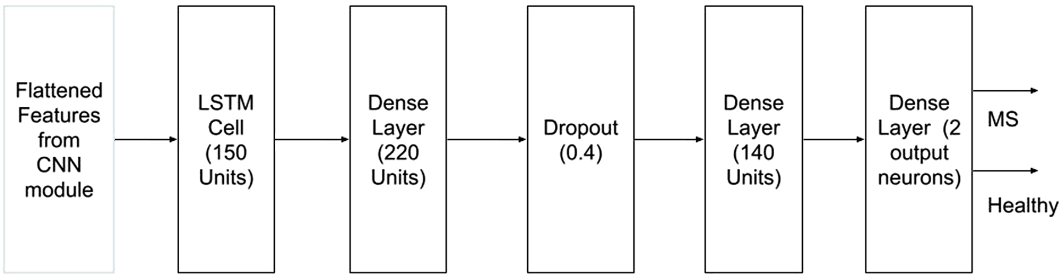
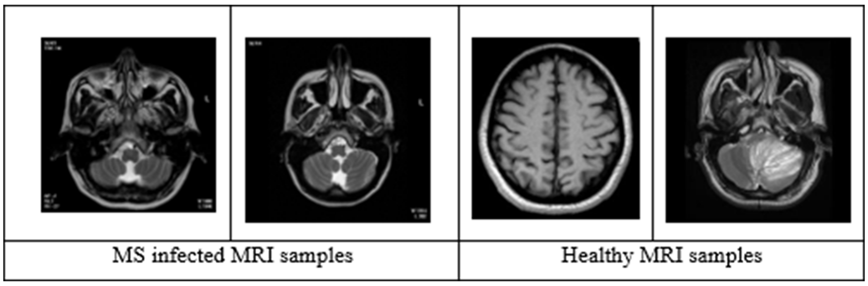


Figure 6
Sample test images from EHealth Lab & IXI dataset



5. Results and Discussion

In this section, we provide an analysis of performance improvement in MS detection accuracy with the proposed hybrid CNN-RNN model. This improvement is achieved by stacking up the features extracted by the two transfer learning models and feeding them to the RNN model for final classification.

5.1. Dataset

5.1.1. EHealth

This dataset consists of MRI scans from 38 patients, comprising 17 males and 21 females, diagnosed with CIS of MS and presenting with MRI-detected brain lesions. The patients have a mean age of 34.1 years, with a standard deviation of 10.5 years. Each patient's MRI data were captured twice, with intervals of 6 to 12 months between the scans. The MRI scans were acquired using a 1.5 Tesla protocol to ensure consistency in imaging quality and resolution. All the images were in .TIFF and .PLQ format. This repeated imaging aids in observing lesion progression and evaluating the course of MS in each patient.

5.1.2. IXI dataset

This dataset involved the collection of approximately 600 MR images from normal, healthy subjects, providing a robust dataset for analysis. Each subject underwent an extensive MR imaging protocol, which included T1, T2, and PD-weighted images, along with magnetic resonance angiography and diffusion-weighted images captured in 15 directions. The data was gathered across three hospitals in London, ensuring diversity in acquisition environments and equipment. Hammersmith Hospital contributed images using a Philips 3T system, with scanner parameters documented for reference. Guy's Hospital and the Institute of Psychiatry used Philips 1.5T and GE 1.5T systems, respectively. The sample MRI samples used for research study are depicted in Figure 6.

5.2. Experimental setup

All the models are trained over the desktop with core i3 10th generation CPU and 8GB RAM memory. The implementation is done in Python version 3.8.3 in Google Colab Notebook environment. The models are constructed using the Keras API, which is imported from TensorFlow. All the models are trained using the hyperparameters listed in Table 1.

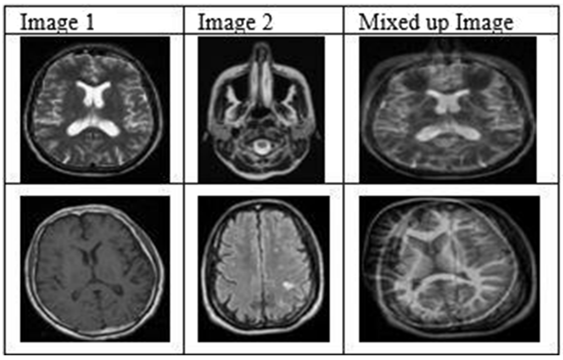
5.3. Data pre-processing & data splitting

To further enhance the generalization and robustness of the proposed hybrid model, a mix-up data augmentation procedure is applied. This step involves randomly selecting a pair of images from the training set and mixing them by assigning weights to each of them based on the alpha value generated by the beta generation

Table 1
Hyper-parameters of the proposed model

Activation function	ReLu, Softmax
Loss function	Binary cross-entropy
Optimizer	Adam
Learning rate	1×10^{-6}
Batch size	32
Number of epochs	25

Figure 7
Mix-up data augmentation on training samples showing variations in texture



technique. Figure 6 depicts the samples of MRI images used for a research study, showcasing the diversity and quality of the images.

By incorporating these data pre-processing steps, the dataset becomes most suitable for training the proposed hybrid CNN-RNN model as it promotes consistency and increases the diversity of training samples. Through the mix-up data augmentation process, approximately 4 additional images are generated for each pair of images from the training samples. Figure 7 shows the visual illustration of mix-up augmentation applied for a pair of images selected from the training set.

The pre-processed data are then split for training, validation, and testing in the ratio 75:10:15, respectively. The amount of image samples for training, validation, and testing is shown in Table 2.

Table 2
Dataset used for training the proposed model

Category	Training	Validation	Testing
MS	1578	210	316
Healthy	1578	210	316

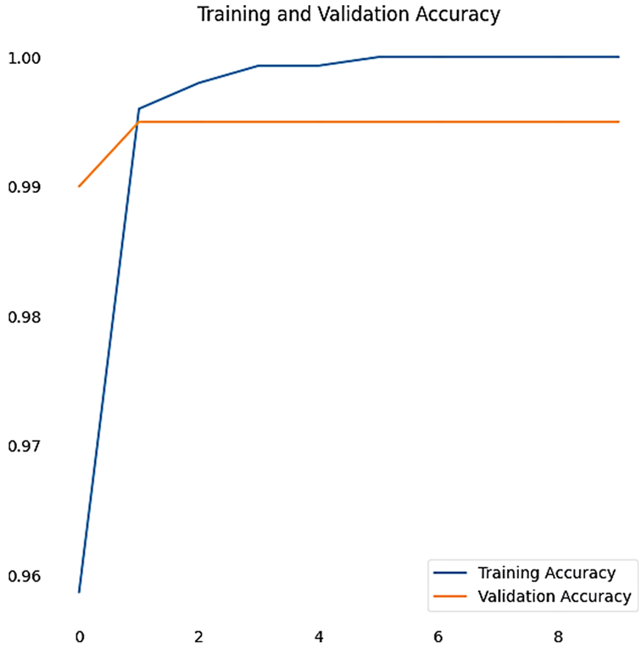
5.4. Feature extraction using transfer learning models

In the first phase of the proposed hybrid model for MS detection, two transfer learning CNN models are built. m-Inception V3 model is constructed by adding custom layers on top of the network after removing the classification layer. This modified architecture extracts a total of 12,800 features. Additionally, a m-DenseNet 121 model is built using transfer learning, leveraging knowledge gained from the ImageNet database. Custom layers are added to the top of this pre-trained model, resulting in the extraction of approximately 25,088 features. The features extracted by both transfer learning CNN models are then stacked up using a feature concatenation process. This approach aims to create a more informative representation of the captured information, which has demonstrated improved results in enhancing the model's performance.

5.5. Feature classification using RNN model

The stacked features are flattened and then passed to the LSTM model, which includes an LSTM layer with 100 units. These LSTM

Figure 8
Training and validation accuracy



units are responsible for capturing temporal dependencies within the input sequences derived from the two CNN models. The model is trained for 10 epochs with a batch size of 32. We utilize sparse categorical cross-entropy as the loss function and the Adam optimizer to adjust the learning rate, initially set to 0.0001. The training and validation accuracy achieved by the proposed model is depicted in Figure 8. The training and validation loss achieved by the proposed model is depicted in Figure 9. The model achieved an impressive score of 99.67% accuracy on the test data.

5.6. Ablation study on varying data augmentation procedure

We conducted an ablation study where we evaluated the performance of our proposed hybrid CNN-RNN model using pre-processed MRI samples alone, as well as in combination with mix-up data augmented MRI samples. Table 3 shows the confusion metric of the proposed model with pre-processed MRI samples. Type-I (Healthy misclassified as MS) and Type–II (MS misclassified as Healthy) errors are 17.18% and 4.13%, respectively. Table 4 shows the confusion metric of the proposed hybrid CNN-RNN model where both MRI samples and mix-up augmented MRI samples are used for training the model.

The measure indicates the model's performance in classifying cases as "Healthy" or "MS". Out of total cases, 149 were correctly predicted as "Healthy" with a rate of 99.33%, and all 150 were correctly classified as "MS" with a rate of 100%. There is one case of "Healthy" that is misclassified as "MS" (false positive) constituting a 0.66% error rate. There is no false negative observation by the model (MS predicted as healthy) resulting in a 0% error rate. These values are critical for assessing the diagnostic accuracy and precision of the model, showing a high rate of correct classifications and a low rate of errors. The performance measures of the proposed model for performing an ablation study are listed in Table 5.

Figure 9
Training and validation loss

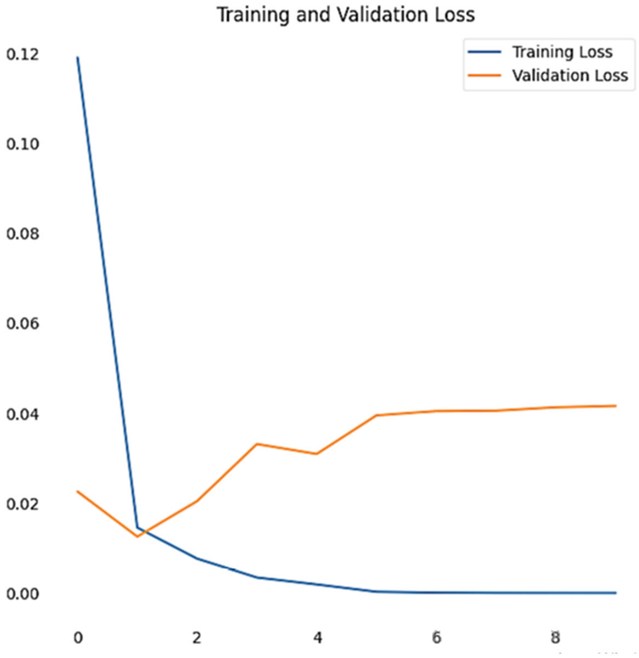


Table 3
Confusion matrix (without mix-up data augmentation)

	Healthy	MS
Healthy	42.6%	7.3%
MS	2.1.%	48.0%

Table 4
Confusion matrix (with mix-up data augmentation)

	Healthy	MS
Healthy	42.6%	0.4%
MS	0.0%	48.0%

Type-I and Type-II errors are reduced when mix-up augmented data is used along with pre-processed data for training the proposed model. This approach enhances the model's generalizability and helps prevent overfitting. From the results, it is evident that our proposed hybrid CNN-RNN model achieves better performance in classifying MS samples as "MS" or "Healthy". We have selected two pre-trained models InceptionV3 and DenseNet121 for extracting features from the MRI modalities. To further enhance the performance of the model, we incorporated LSTM components into our architecture. LSTMs are renowned for their ability to capture intricate temporal dependencies within data, making them exceptionally valuable in medical imaging. By integrating both custom CNN and LSTM components, our hybrid model strives to elevate the accuracy and resilience of medical image analysis. This combination makes use of the strengths of individual architecture, culminating in a robust framework capable of extracting intricate features and discerning complex patterns inherent in medical images.

5.7. Comparison with related works

We evaluated the performance of our proposed hybrid CNN-RNN model for MS detection with related research works. To evaluate performance, we use a comprehensive set of metrics including accuracy, precision, recall, $F1$ score, and the AUC-ROC curve. Table 6 presents a comparative analysis of the performance metrics for the proposed CNN-RNN model with a few state-of-the-art related models for MS detection.

The comparative results confirm that the proposed model outdoes the mentioned state-of-the-art methods achieving better performance measures. We have conducted an experimental study with and without applying a mix-up data augmentation procedure and got an effective classification performance of 99.67% on applying augmentation. Amalgaming CNN with the RNN model has shown improved performance in MS detection.

The proposed model extracts the features from two transfer learning employed CNN models which are then fed into the RNN model to predict MS detection. CNN is good at capturing spatial features of the image, whereas RNN adapts to capture temporal features of the image. The proposed hybrid model leverages the strength of both CNN and RNN models and achieves good performance in MS detection from MRI samples.

Table 5
Performance measure of the proposed model

DL Model	Accuracy	Precision	Recall	AUC
Hybrid CNN-RNN model (without mix-up augmentation)	0.9871	0.9843	0.9754	0.9816
Hybrid CNN-RNN model (with mix-up augmentation)	0.9966	0.9918	0.9846	0.9902

Table 6
Comparison with state-of-the-art methods

Reference	Model	Accuracy
Ekmekyapar and Tasci [14]	Feature Extraction using Mobile NetV2 with exemplar-based learning & classification using KNN classifier	98.42
Ambayiram and Ganesan [15]	White matter segmentation using Chaotic Leader-Selective Particle Swarm Optimization & classification using a hybrid supervised model	98.36
Proposed Hybrid CNN-RNN model	Feature Extraction using two transfer learning applied CNN models & classification using the LSTM model	99.67

However, there are limitations to consider. The proposed model is trained and tested on the limited dataset. To improve the generalizability of the system, the model has to be trained on multiple diverse datasets. Interpretability of the proposed model's decision is a big challenge, and efforts should be made to develop an explainable AI model to improve the trustworthiness of AI involved in clinical workflow by medical practitioners.

This digital replica created by analyzing medical images using a deep learning model behaves as a DT of patients with MS. This allows clinicians to anticipate potential complications in patients sharing similar health conditions to the DT. Consequently, clinicians can take proactive measures to plan treatments and prevent health complications.

6. Conclusion

In this research, an HDT framework is proposed for detecting MS. The model comprises two main components: feature extraction and feature fusion for classification. Modified versions of InceptionV3 and DenseNet121 models are utilized to extract features from MRI samples, which are then stacked up and fed into the LSTM RNN model to capture temporal dependencies within MRI modalities. Dropout layers are included to prevent overfitting in the LSTM layer. This hybrid approach outperforms existing models, achieving an accuracy of 99.67%. This model will aid the Clinicians to detect MS and plan treatments to improve their well-being. Future work will involve integrating MRI data with patient disability scores to predict disease severity, with the ultimate goal of developing a web-based tool for clinicians to offer personalized treatment and rehabilitation recommendations to MS patients.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data availability statement

The data that support the findings of this study are openly available in IXI Dataset at https://brain-development.org/ixi-dataset/.

Author Contribution Statement

Ramya Palaniappan: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing – original draft, Writing – review & editing. **Siva Rathinavelayutham:** Investigation, Resources, Data curation, Visualization, Supervision, Project administration.

References

[1] Jacob, A., Shatila, A. O., Inshasi, J., Massouh, J., Mir, R., Noori, S., & Yamout, B. (2024). Disease modifying treatment guidelines for multiple sclerosis in the United Arab Emirates. *Multiple Sclerosis and Related Disorders*, *88*, 105703. https://doi.org/10.1016/j.msard.2024.105703

[2] Shoeibi, A., Khodatars, M., Jafari, M., Moridian, P., Rezaei, M., Alizadehsani, R., . . . , & Acharya, U. R. (2021). Applications of deep learning techniques for automated multiple sclerosis detection using magnetic resonance imaging: A review. *Computers in Biology and Medicine*, *136*, 104697. https://doi.org/10.1016/j.compbiomed.2021.104697

[3] Subramanian, K. (2020). Digital twin for drug discovery and development—The virtual liver. *Journal of Indian Institute of Science*, *100*, 653–662.

[4] Coorey, G., Figtree, G. A., Fletcher, D. F., Snelson, V. J., Vernon, S. T., Winlaw, D., . . . , & Redfern, J. (2022).The health digital twin to tackle cardiovascular disease—A review of an emerging interdisciplinary field. *NPJ Digital Medicine*, *5*, 126.

[5] Liu, Y., Zhang, L., Yang, Y., Zhou, L., Ren, L., Wang, F., . . . , & Deen, M. J. (2019). A novel cloud-based framework for the elderly healthcare services using digital twin. *IEEE Access*, *7*, 49088–49101.

[6] Laamarti, F., Badawi, H. F., Ding, Y., Arafsha, F., Hafidh, B., & Saddik, A. E. (2020). An ISO/IEEE 11073 standardized digital twin framework for health and well-being in smart cities. *IEEE Access*, *8*, 105950–105961.

[7] Hussain, I., Hossain, M. A., & Park, S. J. (2021). A healthcare digital twin for diagnosis of stroke. In *IEEE International Conference on Biomedical Engineering, Computer and Information Technology for Health*, 18–21.

[8] Sasikaladevi, N., & Revathi, A. (2024). Digital twin of renal system with CT-radiography for the early diagnosis of chronic kidney diseases. *Biomedical Signal Processing and Control*, *88*(Part A), 105632. https://doi.org/10.1016/j.bspc.2023.105632

[9] Ojo, S., Krichen, M., Alamro, M. A., Mihoub, A., Avelino Sampedro, G., & Kniezova, J. (2024). Improving multiple sclerosis disease prediction using hybrid deep learning model. *Computers, Materials and Continua*, *81*(1), 643–661. https://doi.org/10.32604/cmc.2024.052147

[10] Zhang, Y. D., Pan, C., Sun, J., & Tang, C. (2018). Multiple sclerosis identification by convolutional neural network with dropout and parametric ReLU. *Journal of Computational Science*, *28*, 1–10. https://doi.org/10.1016/j.jocs.2018.07.003

[11] Özkaraca, O., Bağrıaçık, O. İ., Gürüler, H., Khan, F., Hussain, J., Khan, J., & Laila, U. E. (2023). Multiple brain tumor classification with dense CNN architecture using brain MRI images. *Reproductive and Developmental Biology*, *13*(2), 349. https://doi.org/10.3390/life13020349

[12] Jain, S., Rajpal, N., & Yadav, J. (2020). Multiple sclerosis identification based on ensemble machine learning technique. *Social Science Research Network*. https://doi.org/10.2139/SSRN.3734806

[13] Öztürk, Ş., & Özkaya, U. (2020). Gastrointestinal tract classification using improved LSTM based CNN. *Multimedia Tools and Applications*, *79*, 28825–28840. https://doi.org/10.1007/s11042-020-09468-3

[14] Ekmekyapar, T., & Taşcı, B. (2023). Exemplar mobileNetV2-based artificial intelligence for robust and accurate diagnosis of multiple sclerosis. *Diagnostics*, *13*, 3030. https://doi.org/10.3390/diagnostics13193030

[15] Pandian, A., & Udhayakumar, G. (2023). Improved multiple sclerosis diagnosis with advanced deep learning techniques. *International Journal of Imaging Systems and Technology*, *33*, 2128–2141. https://doi.org/10.1002/ima.22908

[16] Zhang, K., Zhou, H. Y., Baptista-Hon, D. T., Gao, Y., Liu, X., Oermann, E., . . . , & Wu, J. (2024). Concepts and applications of digital twins in healthcare and medicine. *Patterns*, *5*(8), 101028. https://doi.org/10.1016/j.patter.2024.101028

[17] Awasthi, M., Raghuvanshi, C. S., Dudhagara, C., & Awasthi, A. (2025). Chapter 17: Exploring virtual smart healthcare trends using digital twins. In S. Iyer, A. Nayyar, A. Paul & M. Naved (Eds.),

Digital twins for smart cities and villages (pp. 377–406). Elsevier. https://doi.org/10.1016/B978-0-443-28884-5.00017-8

[18] Derraz, B., Breda, G., Kaempf, C., Baenke, F., Cotte, F., Reiche, K., . . . , & Gilbert, S. (2024). New regulatory thinking is needed for AI-based personalized drug and cell therapies in precision oncology. *npj Precision Oncology*, *8*, 23. https://doi.org/10.1038/s41698-024-00517-w

[19] Angulo, C., Gonzalez-Abril, L., Raya, C., & Ortega, J. A. (2020). A proposal to evolving towards digital twins in healthcare. In *International Work-Conference on Bioinformatics and Biomedical Engineering Granada Spain*, 418–426.

[20] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., . . . , & Rabinovich, A. (2014). Going deeper with convolutions. *arXiv Preprint: 1409.4842*.

[21] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2018). Densely connected convolutional networks. *arXiv Preprint: 1608.06993*.

How to Cite: Palaniappan, R., & Rathinavelayutham, S. (2025). Digital Twin-Assisted Deep Learning with Model Fusion for Detecting Multiple Sclerosis in MRI Modalities. *Journal of Data Science and Intelligent Systems*. https://doi.org/10.47852/bonviewJDSIS52024391

Appendix

Algorithm 1: Hybrid CNN-RNN model for MS detection

Input: Training Dataset D = {(x1, y1), (x2, y2) ..(xn, yn)}
First-level feature extractors: C1, C2 . . . . Ck
Meta-level classifier: Cˆ
Output: Trained Ensembled Classifier Mˆ

Begin
Step-1: Train first-level classifiers on training dataset D
for i = 1,2 . . . . k
Fi = Ci(D) % Extract features
end for

Step 2: Construct a new meta-dataset from predictions
extracted from D
Dˆ= F1 + F2 + ... + F_k % Stack up extracted features

Step 3: Train meta-level classifier on a new dataset
Mˆ= Cˆ (Dˆ) % Classification
Return Mˆ

End