IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 16 July 2024, accepted 5 August 2024, date of publication 8 August 2024, date of current version 19 August 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3440318

RESEARCH ARTICLE

Deep Attention V-Net Architecture for Enhanced Multiple Sclerosis Segmentation

V. P. NASHEEDA[1] AND VIJAYARAJAN RAJANGAM[2], (Senior Member, IEEE)
[1]School of Electronics Engineering, Vellore Institute of Technology, Chennai 600127, India
[2]Centre for Healthcare Advancement, Innovation, and Research, Vellore Institute of Technology, Chennai 600127, India

Corresponding author: Vijayarajan Rajangam (viraj2k@gmail.com)

This work was supported by the Vellore Institute of Technology, Chennai, India.

ABSTRACT The central nervous system is affected by multiple sclerosis (MS) which destroys the neurocommunication. Among the diagnostic imaging systems, magnetic resonance imaging is the most preferred one to track new and enlarged MS lesions. In this paper, we propose a deep-attention V-Net architecture with modified compression and expansion sections to segment MS. The first network performs feature extraction and expansion, thus delivering enhanced feature maps for segmenting the region of interest. The second network performs feature extraction with modified V-Net architecture and performs segmentation using the soft-max function. This model is evaluated on the publicly available MICCAI 16, MSSEG-2, and Brain MRI Dataset of Multiple Sclerosis with Consensus Manual Lesion Segmentation and Patient Meta Information dataset (2022) datasets and compared with the existing models. The proposed deep-attention V-Net model is also compared with sequential models, using V-Net and U-Net in terms of precision, sensitivity, accuracy, loss, mean IOU, F1 Score, and dice score. The suggested approach delivers a dice score of 0.8900 using the MICCAI 16 dataset, 0.9000 using the MSSEG-2 dataset and 0.9638 using the combined MSSEG 2 and Brain MRI Dataset of Multiple Sclerosis with Consensus Manual Lesion Segmentation and Patient Meta Information dataset (2022) datasets. These dice score values are superior to other deep-learning networks.

INDEX TERMS CNN, deep learning, lesion, MRI, MICCAI, multiple sclerosis.

I. INTRODUCTION

Multiple Sclerosis (MS) is a progressively demyelinating and neurodegenerative autoimmune disease characterized pathologically by mononuclear inflammatory cells, axonal loss, gliosis with the focal formation and diffused abnormalities in the central nervous system (CNS) [1]. The lesions or plaques are formed due to multiple sclerosis in the CNS. These lesions are commonly circular to elliptical and vary in diameter from millimeters to one-centimeter [1]. The initial stage of MS has separate lesions. As the MS advances, these lesions will join together. Normally MS lesions form in periventricular and juxta cortical white matter [1]. The indications of MS are optic disorder, aches, weakness, and imperfect motor functions. The signs and span are different for different persons. MS is affected by age, gender, family history, climate, etc [2]. Among the imaging techniques suggested for MS diagnosis, Magnetic Resonance Imaging (MRI) is a non-penetrating visual representation method that is widely preferred for brain imaging. The different modes of MRI and the appearance of brain tissues [3] are presented in Figure 1. Identification and segmentation of MS lesions play crucial roles in aiding MS diagnosis and follow-up. The human-assisted segmentation of MS lesions is tedious and liable to intra and inter-reader changes [4]. The automatic segmentation of MS lesions is complex due to the lack of information regarding the size, shape, and localization. Several automatic segmentation models have been proposed in the last two decades for MS segmentation [2]. Machine Learning (ML) and Deep Learning (DL) methods have been proposed to observe the fundamental and hidden attributes of MRI data. The notable advantage of DL-based methods is the automatic feature extraction that leads to effective segmentation [5]. This paper proposes a DL model named Deep Attention V-Net (DAV_{Net}) for MS lesion segmentation. This model has two networks. A Deep

The associate editor coordinating the review of this manuscript and approving it for publication was Li Zhang.

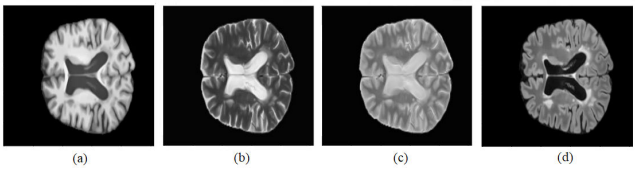

FIGURE 1. Different types of MRI. (a) T1-w MRI, (b) T2-w MRI,(c) PD MRI and (d) FLAIR MRI of an MS patient.

Attention Network (DA_{Net}) for giving attention to the MS tissues, and a Modified V-Net (MV_{Net}) architecture for MS segmentation. Each network has compression and expansion paths. Every path has convolutional layers, non-linearity function, and skip connections. One more block with the above-said components is introduced in each network to provide more attention for the extraction of the region of interest for improved segmentation. The effectiveness of the suggested network is analyzed by accuracy, mean IOU, precision, sensitivity, specificity, loss, and dice score. The important findings of our work are itemized below.

- DA_{Net} which focuses on the MS tissues leading to improved segmentation followed by a MV_{Net} architecture for MS segmentation.
- Implemented Other sequential networks like V-Net, U-Net: MV_{Net}, and MV_{Net}: U-Net and proposed model on MICCAI 16 dataset, This experiment proves the effectiveness of the proposed sequential model.
- The proposed method delivers good results for smaller and more complex MS lesions with better dice scores, F1 scores, accuracy, mean IOU, loss, precision, sensitivity, and specificity.

This paper is systematized as follows, Section II is about the related works and publicly available MS datasets for MS lesion segmentation. Section III is about the proposed method it includes the network architecture and a detailed description of all the blocks. Section IV presents results and discussion including dataset description, performance measures, and eight conducted experiments followed by a conclusion in Section V.

II. RELATED WORK AND MS DATASETS

This section presents various DL models proposed for MS lesion segmentation followed by publicly available different MS Datasets.

A. DL NETWORKS FOR MS LESION SEGMENTATION

There are many DL models for MS segmentation and classification. It includes Convolutional Neural Network (CNN), U-Net, Fully Convolutional Neural Networks (FCNN), Generative Adversarial Networks (GAN), Recurrent Neural Networks (RNN), Residual Network (ResNet), Inception V3-Net, etc. The CNN architecture comprises a convolutional layer, pooling, batch normalization, and fully connected layers [2]. U-Net is a CNN with two sections such as encoding and decoding sections. The encoder has several down-sampling segments together with the convolutional layers and the decoder has several up-sampling sections and layers of convolution [6]. The increase in lesion size results in robust segmentation and improvement of segmentation [7]. To get the best outcome from CNN, the number of layers can be increased.

In the inception network [8], wider layers are considered instead of considering deeper layers. The inception V3-Net has three various convolutions along with one max pooling layer. The main feature of inception V3-Net is the splitting of the convolutional kernel which will make the model flexible for different scales and also reduce overfitting. Another DL model for reducing the vanishing gradient problem is the Residual Network (ResNet). It has mainly two blocks such as identity and convolutional blocks. In this network, every residual layer undergoes convolution with the inputs of the upcoming layer. The shortcut connections are present to add the inputs and outputs [9].

CNN-based DL architectures are used for 2D medical image segmentation in an efficient way [10], [11] [12]. With the availability of three-dimensional (3D) imaging data for disease prediction, CNN-based 3D segmentation is also available. Vaidya et al. [13] proposed a 3D 4-layer CNN-based method in 2015. All the CNN layers use the softmax activation function. In 2017, Valvendra et al. [14] widened the research to use 3D information for the segmentation. Millertari et al. [15] paper is based on CNN with the hough voting approach but this is applicable only for blob-like shapes and not an end-to-end phenomenon. Another method is a 3D segmentation by Kholmovski et al. [12], which is an all-along-in-line method but not a deep network structure on multiple scales.

Isensee et al. [16] suggested a DL model called nnu-net for image segmentation which uses the encoder-decoder hyperparameters to specify the dataset. It uses adaptive pre-processing, training schemes, and inference. A work by Brosch et al. [17] is a seven-layer network based on a 3D convolutional encoder with a shortcut connection for the segmentation of lesions with different sizes. SMORE is a method by Zhao et al. [18] based on CNN. It is an end-to-end learning network and it puts back the image quality by increasing the resolution and reducing the aliasing. Kang, et al. [19] proposed an end-to-end network called Attention Context U-Net (ACU-net) for MS lesion segmentation. it has an attention block that will enrich the spatial details and attribute depiction in an extraction step.

Raura et al. [20] worked on that statistical evaluation which gave a better performance concerning precision while keeping similar results on sensitivity and Dice similarity. Schmidt et al. [21] proposed a threshold-based lesion detection method. The 3D-based image segmentation mainly focused on features and labels for the prediction.

The summary and the performance measures of different DL-based models for MS lesion segmentation are

TABLE 1. Related Works.

Reference	Details of network layers	Datasets used	Evaluation measures
Vaidya et al. [13]	Layer1(L1) 60 filters of 4x4x4 with average pooling of 2x2x2.L2 60 filters of 3x3x3 with average pooling of 2x2x2. L3: Multi-layer perception L4: Softmax	21 Clinical images	Dice score=0.833,sensitivity=0.876
Valverde et al. [14]	4 Convolutional, 2 Max-pooling, 3 fully connected layers, 3 drop out layers	ISBI 2015	F1 Score=0.63, SEN=0.55, Precision =0.79.
Brosche et al. [17]	Convolutional, average-pooling, convolutional, deconvolutional, up-pooling, deconvolutional layers	ISBI 2015	F1 Score=0.638
McKmey et al. [23]	17 Convolutional, 16 Batch normalization, 3 max pooling, 3-Up sampling, concatenation	MICCAI 2016	
Hashemi Seyed Raein et al. [24]	5 Convolutional, 3 batch normalization, 11 Dense blocks, 5 transition down, 5 transition up, Deconvolution, 5 concatenation	MICCAI 2016	F1 Score=0.699
Ghosal et al. [25]	10 Convolutional, 8 B batch normalization, 4 max pooling, 3-Up sampling.	MICCAI 2016	F1 Score=0.76, Accuracy= 0.9679, Sensitivity=0.65, Specificity= 0.86
Afsal et al. [26]	6 Convolutional, 6 max-pooling layers	ISBI 2015	F1 Score=0.67, Sensitivity=0.48
McKmey et al. [27]	2 Convolutional, 4 Dense blocks, max-pooling, Up sampling	MICCAI 2016	F1 Score=0.60, S1-score=0.57

presented in Table 1. To enhance the performance measures of MS lesion segmentation, we put forward the DAV_{Net} architecture.

B. MS DATASETS

The problem of smaller size datasets for MS detection and classification was somehow compensated by the Medical Image Computing and Computer-Assisted Intervention (MICCAI) and IEEE International Symposium Imaging (ISBI) MS lesion segmentation challenge datasets. The description of MS datasets used for many studies is given in Table 2. MICCAI 2016 challenge was organized for the study of multiple sclerosis segmentation algorithms. It includes an extensive range of automatic algorithms for independent evaluation comprising thirteen methods for segmenting MS lesions against fifty-three MS cases. This dataset contains T1-w, T2-w, FLAIR, and PD-w images and was acquired through various MRI scanners using different magnetic field strengths [22]. The MICCAI 2021(MSSEG-2) dataset has a total of 100 images. The brain MRI Dataset of multiple sclerosis with Consensus Manual Lesion Segmentation and Patient Meta Information-2022 (Brain MRI for MS-2022) dataset has 60 patient data. The proposed network is tested on the MICCAI 2016 dataset for comparison with the existing models and four sequential DL networks. The proposed network was also trained and tested on the MSSEG-2 dataset. For generalization, the network is also trained on the MSSEG-2 dataset and tested on brain MRI for the MS- 2022 dataset.

C. PRE-PROCESSING

Pre-processing is applied to the raw dataset before the segmentation process. These steps are applied to improve the data quality, noise reduction, artifact correction, and make the data more suitable for segmentation. In this work, the denoising is performed by median filtering and intensity normalization by histogram matching which reduces the intensity variation due to different factors like scanner setting, patient positioning, etc. The resampling was also done to standardize the dimensions of the data. Here the downsampling was applied to each MRI data of the MICCAI dataset with a reduction factor of 8. The downsampling factor of 2 is used for the Brain MRI MS-2022 dataset.

TABLE 2. MS Datasets.

Dataset	Modalities	No. of cases	MRI scan
MICCAI 2008 [28]	T1-w, T2-w and FLAIR	45	3T Seimens Allegra, 3T seimens
ISBI 2015 [11]	T1-w, T2-w, PD-w and FLAIR	19	3T Philips
MICCAI 2016	T1-w, T2-w and FLAIR	53	Seimen Aera 1.5T, Seimens verio 3T, Philips ignia 3T, General Electric Discovery 3T
MICCAI 2021(MSSEG-2)	FLAIR	100	
Brain MRI for MS-2022	FLAIR	60	1.5 Tesla

TABLE 3. Network description of DAV_{Net}.

Block	Input image size	Number of channels	Output image size
1	$128 \times 128 \times 64$	16	$64 \times 64 \times 32$
2	$64 \times 64 \times 32$	32	$32 \times 32 \times 16$
3	$32 \times 32 \times 16$	64	$16 \times 16 \times 8$
4	$16 \times 16 \times 8$	128	$8 \times 8 \times 4$
5	$8 \times 8 \times 4$	256	$4 \times 4 \times 2$
6	$4 \times 4 \times 2$	512	$8 \times 8 \times 4$
7	$8 \times 8 \times 4$	512	$16 \times 16 \times 8$
8	$16 \times 16 \times 8$	256	$32 \times 32 \times 16$
9	$32 \times 32 \times 16$	128	$64 \times 64 \times 32$
10	$64 \times 64 \times 32$	64	$128 \times 128 \times 64$
11	$128 \times 128 \times 64$	32	$128 \times 128 \times 64$
12	$128 \times 128 \times 64$	16	$64 \times 64 \times 32$
13	$64 \times 64 \times 32$	32	$32 \times 32 \times 16$
14	$32 \times 32 \times 16$	64	$16 \times 16 \times 8$
15	$16 \times 16 \times 8$	128	$8 \times 8 \times 4$
16	$8 \times 8 \times 4$	256	$4 \times 4 \times 2$
17	$4 \times 4 \times 2$	512	$8 \times 8 \times 4$
18	$8 \times 8 \times 4$	512	$16 \times 16 \times 8$
19	$16 \times 16 \times 8$	256	$32 \times 32 \times 16$
20	$32 \times 32 \times 16$	128	$64 \times 64 \times 32$
21	$64 \times 64 \times 32$	64	$128 \times 128 \times 64$
22	$128 \times 128 \times 64$	32	$128 \times 128 \times 64$

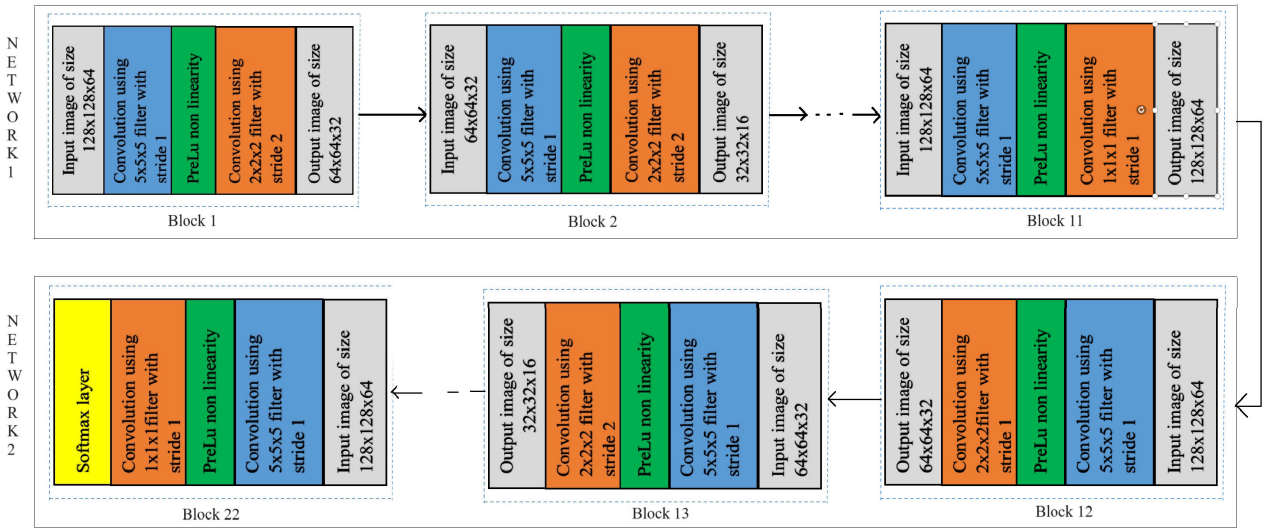

FIGURE 2. Block diagram of DAV_{Net}.

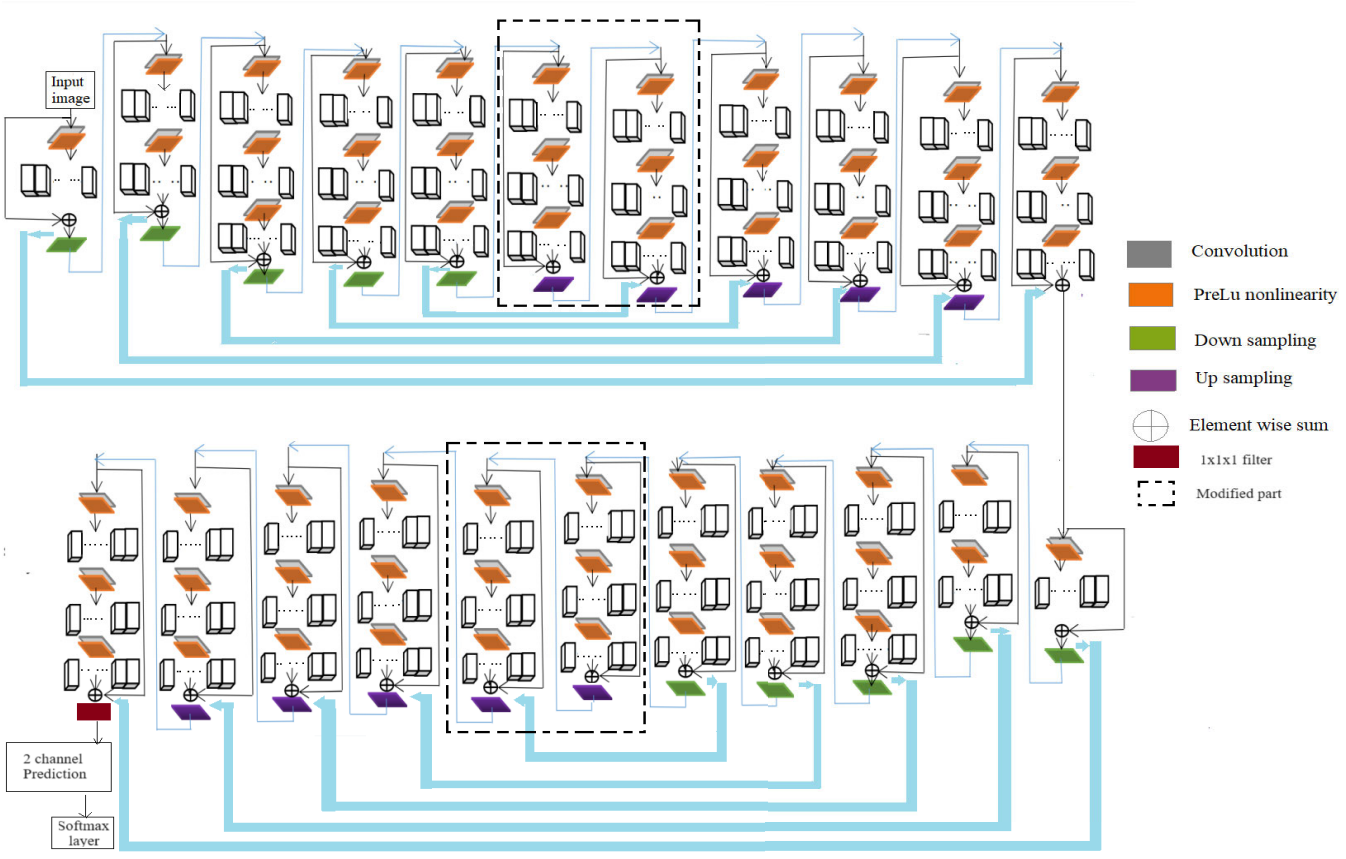

FIGURE 3. DAV_{Net} architecture.

III. PROPOSED METHOD

The DAV_{Net} is a modified sequential V-Net architecture. It does the role of squeezing the features in the form of feature maps for segmentation using encoding and decoding paths. The DAV_{Net} has two networks named as DA_{Net} and MV_{Net}. Both networks have 11 blocks and hence there are

22 blocks in DAV_{Net}. The feature maps from the last block of DA_{Net} are combined and given to the first block of MV_{Net} (block 12). In DA_{Net}, the feature maps that are very much relevant to segmentation are extracted by compression and expansion paths. There is no softmax function in the first network. The block diagram of DAV_{Net} is shown in Figure 2 and the architecture of DAV_{Net} is shown in Figure 3. The Input image was initially given to DA_{Net} and will pass through the different blocks of DA_{Net} then the enhanced region of interest is given to MV_{Net}. The network description of DAV_{Net} is given in Table 3. Each block of DAV_{Net} has a convolutional layer, PReLU nonlinearity, downsampling, or up-sampling layers. The convolutional operation is done by applying the kernel to the input image. The output will be the convoluted feature maps. Each link of the convolutional layer shown in Figure 3 has a unique kernel and produces different feature maps. This convolution is followed by the PReLU nonlinear function for producing nonlinear decision boundaries of the feature map. The downsampling of the feature map is done by the convolutional kernel of size $2 \times 2 \times 2$ with a stride two. Then the resultant image size is cut in half. All the first six blocks perform the same operation with twice the number of feature maps in each block. The 6^{th} block reduces the feature map to $4 \times 4 \times 4$ and the 7^{th} block does the expansion by a factor of two by the deconvolution. Blocks $7-11$ do the deconvolution with the location information from the compression path. At the end of the 11^{th} block, the output image dimension is the same as the input image dimension and supplied to the input of MV_{Net}. The layers of MV_{Net} perform the same operations as DA_{Net}, and softmax activation is applied at the last block of MV_{Net} for the classification of pixels. Each pixel will be classified into MS pixel or non-MS pixel. The convolution kernels present in DA_{Net} perform feature extraction, meanwhile, the expansion path reconstructs the image to the original spatial resolution. Here, the details in the input image are appropriately retained to meet the requirement of segmentation. The feature maps from the DA_{Net} deliver enhanced image details for better segmentation which can further be processed by MV_{Net} for segmentation using the softmax function. The pixel-wise segmentation can be represented by binary numbers 1 to indicate the presence of MS lesion and 0 for the normal tissue. The segmentation can be represented by the following expression [17].

$$\mathrm{A} = \operatorname{argmin} \sum_{n=0}^{N-1} (E(S_n, s(I_n)) \tag{1}$$

where, I_n: Training images, S_n: Segmentation, E: Error that occurred in the segmentation process.

A. NETWORK ARCHITECTURE

DAV_{Net} is a sequential framework consisting of two networks. Each network has compression and expansion paths with eleven blocks. We are applying downsampling to five blocks and upsampling to the next five blocks followed by the softmax activation function to the last block of the second network, MV_{Net}.

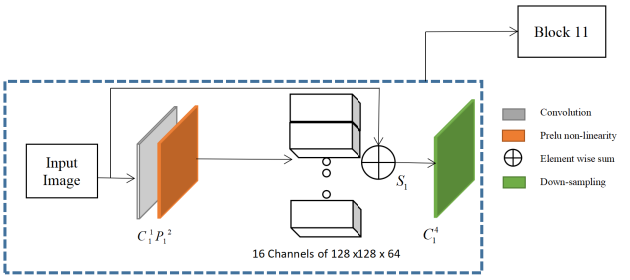

FIGURE 4. Internal diagram of block 1.

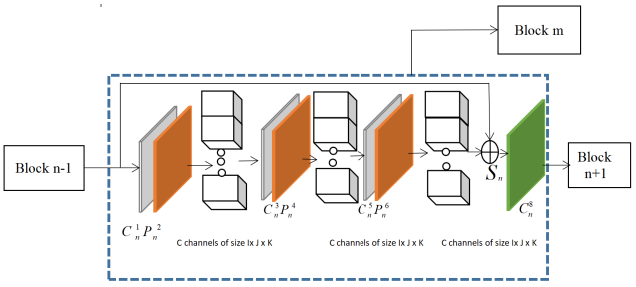

FIGURE 5. Internal diagram of block n.

1) DA_{NET}

DA_{Net} performs feature extraction using compression and expansion paths. The feature maps from block 11 are combined to get a single feature map. Different blocks of DA_{Net} are presented below.

2) BLOCK 1

The internal architecture of block 1 is shown in Figure 4. Each block in the DA_{Net} has a convolutional kernel of size $5 \times 5 \times 5$ with stride 1. This kernel will move on the entire image and provide the features as the output. The input image is initially used for the first convolutional layer of the first path. It is represented as

$$I_j^{(K)} = \mathrm{Max} \sum_{n=0}^{N-1} \overline{W}_{n,i,j}^{(k)} * I_{(i)}^{(k1)} + a_{(j)}^{(k)} \tag{2}$$

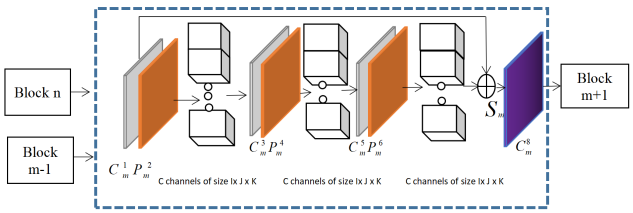

FIGURE 6. Internal diagram of block m.

where, $I_j(K)$ feature map corresponding to the filter ($W_{(i,i,j)}$, k: convolutional layer index, $\alpha_{(j)}$: bias factor. In the DAV_{Net},

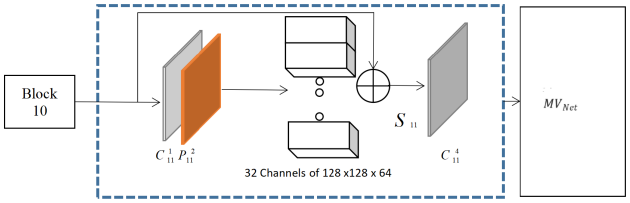

FIGURE 7. Internal diagram of Block 11.

the convolution kernel has a size of 5 × 5 × 5. The term 'j' belongs to (1, M) where M is the number of kernels in the current layer. It is three in each layer of DAV_{Net}. The symbol '*' corresponds to the convolution operation. $\overline{W}$ is the inverted version of W. An activation function, parametric rectified linear unit (PReLU), is given to the outcome of the convolution layer. PReLU can be defined as,

$$f(\bar{I}_j) = \bar{I}_j, \; if \; \bar{I}_j > 0 \tag{3}$$

$$f(\bar{I}_j) = c\bar{I}_j, \; if \; \bar{I}_j < 0 \tag{4}$$

where 'c' is a small and fixed value. This PReLU results in only fewer extra parameters which is equal to the total number of channels. The resultant parameter size is much less relative to the whole number of weights present, hence the risk of overfitting is reduced. Block 1 has 16 channels of size 128 × 128 × 64, a summation block, and a convolutional layer of size 2 × 2 × 2 with a stride 2. This block extracts the features by considering non-overlapping patches. The output of this layer is,

$$I_j^{'(K)} = \text{Max}(0, \sum_{n=0}^{N-1} \overline{Z}_{n,i,j}^{(k)} * f(I_{(j)})^{(k-1)}) + b_{(j)}^{(k)}) \tag{5}$$

where, $\overline{Z}_{n,i,j}$ is a 2 × 2 × 2 filter, $b_{(j)}$: bias factor.

The convolution operation with stride two will reduce the size of a feature map by a factor of two. The number of feature extraction kernels doubles at each block of the compression path such that the size of the image reduces and the feature map size increases. The location information from the compression path will be forwarded to the expansion path.

Then the block 1 output is given to block 2. There will be a reduction of dimension by a factor of 2 while moving from block 1 to block 2.

3) BLOCK 2-6

The block 1 output is given to block 2. The output of block 2 is given to block 3, and so on. Figure 5 shows the internal block diagram of block n where n is the number of blocks $n = 1, 2 \ldots 6$. The channels are illustrated by $i \times j \times k$. For the blocks $n = 2, 3, 4, 5$, and 6, the input image size and the number of channels are presented in Table 4. Every block comprises a convolutional layer of size 5 × 5 × 5 with a stride one, PReLU, summation block, and a convolutional layer of size 2 × 2 × 2 with a stride two. The size of the image becomes halved after each block. The line indicates the passage of location information to block m = 10,9,8,7 respectively from n = 2,3,4, and 5 for the expansion process.

Algorithm 1 DAV_{Net}

- **Input:** Brain MRI along with MS masks.
- **Output:** Segmented images.

Step 1: Select MRI of n patients N_i with the corresponding masks M_i where i=1,2,..n.
Step 2: Change the size of the Image.
Step 3: Divide the images in to n_{T1} images for training, n_{T2} for validation, and n_{T3} for testing.
Step 4: Construct the DAV_{Net} model, S with convolutional layers and PReLU layers as shown in figure 3.

- Applying convolutional mask on inputs
  $\text{I}_j^{(K)} = \text{Max} \sum_{n=0}^{N-1} \overline{W}_{n,i,j}^{(k)} * I_{(i)}^{(k1)}) + a_{(j)}^{(k)}$.
- PReLU Nonlinearity
  $\text{f}(\bar{I}_j) = \begin{cases} \bar{I}_j, if \bar{I}_j > 0 \\ c\bar{I}_j, if \; \bar{I}_j < 0 \end{cases}$
- Feature extraction by nonoverlapping patches
  $I_j^{'(K)} = \text{Max}(0, \sum_{n=0}^{N-1} \overline{Z}_{n,i,j}^{(k)} * f(I_{(j)})^{(k-1)}) + b_{(j)}^{(k)})$
- Deconvolution Operation
  $\bar{I}_j^{(K-1)} = \text{Max}(0, \sum_{n=0}^{N-1} \overline{W}_{n,i,j}^{(k)} * I_{(i)}^{(k-1)}) + a_{(j)}^{(k-1)})$
- Predicted output: Output of softmax layer

Step 5: Train the model S using n_{T1} images with its MS mask to obtain the trained model S_1.training steps are given below.

- Initialize the model with random weights.
- Select the binary cross entropy as a loss function.
- Select Adam optimizer with a learning rate of 0.001 to minimize the loss function.
- Repeat the process of forward propagation, backward propagation, and evaluation iteratively until the model doesn't improve significantly (early stopping criteria).

Step 6: Validate the trained model S_1 with n_{T2} images with its MS mask.The step should be performed iteratively until the early stopping.
Step 7: Evaluate the performance of the the model S_1 with n_{T3} images.

4) BLOCK 7-11

The spatial expansion of lower-resolution images is done with the help of the feature map derived from the compression

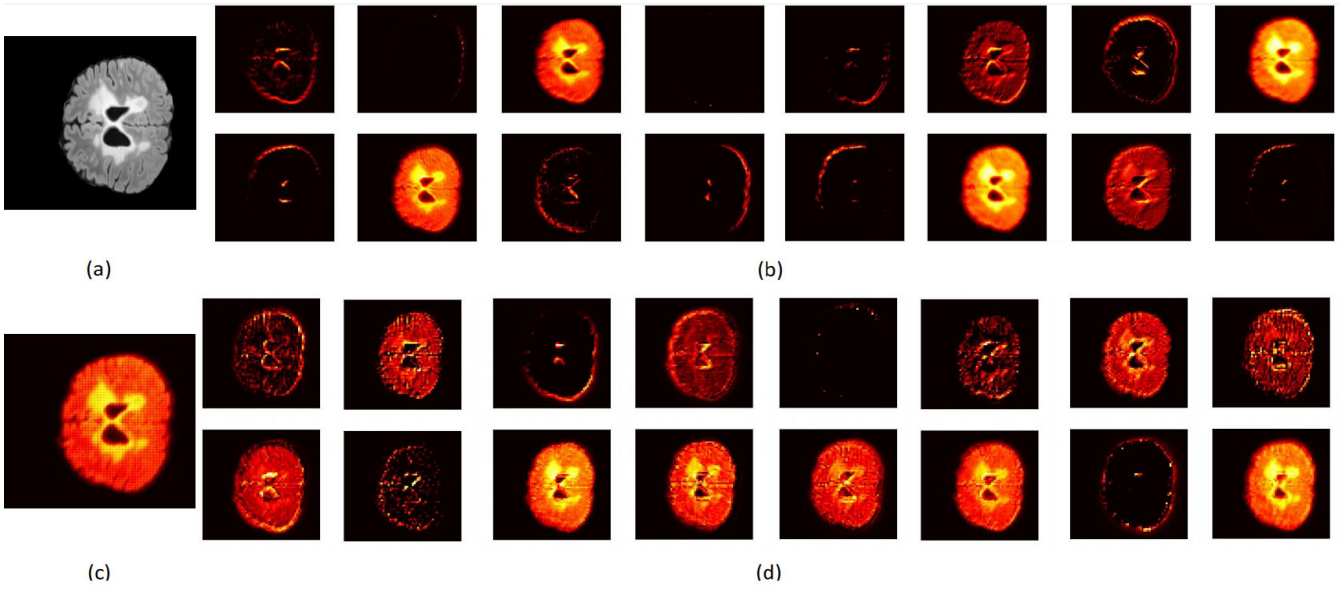

FIGURE 8. (a) Input image (b) Heat maps of the feature maps from the output of Block 11 (c) Heat map of input to block 12 (d) Heat maps of the feature maps from the output of Block 12.

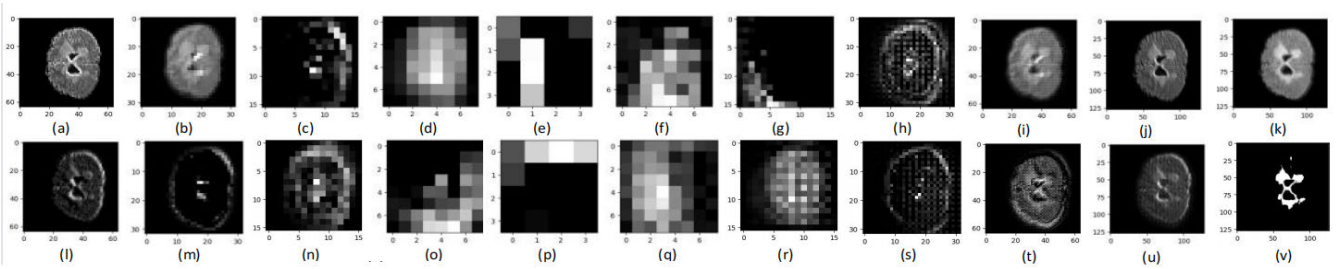

FIGURE 9. Outputs from different blocks. (a)-(k) Output from blocks 1-11 of DA_{Net} respectively. (l)-(v) Output from blocks 12-22 of MV_{Net} respectively.

block. Figure 6 shows the internal block diagram of block m. Where m is the block number of the expansion process. The value of m is 7,8,9,10 corresponding to the blocks 7,8,9 and 10 and channel sizes of $8 \times 8 \times 4$, $16 \times 16 \times 8$, $32 \times 32 \times 16$, $64 \times 64 \times 32$, $128 \times 128 \times 64$ respectively. These blocks use the location information from corresponding blocks for the expansion process. The up convolution will increase the image size by a factor of 2.

Figure 7 shows the internal block diagram of block 11. It has 32 channels of $128 \times 128 \times 64$ and up convolution with $2 \times 2 \times 2$ kernel with a stride 2 for the expansion process. Then, the output of block 11 is given to MV_{Net}.

Block 11 does the expansion operation. it includes deconvolution operations. It is done with $\overline{W}_{(n,i,j)}$

$$\bar{I}_j^{(K-1)} = \text{Max}(0, \sum_{n=0}^{N-1} \overline{W}_{n,i,j}^{(k)} * I_{(i)}^{(k-1)}) + a_{(j)}^{(k-1)}) \quad (6)$$

It unfolds the spatial support of lower pixel density details by combining the important information. The size interpolation is done by deconvolution with a $2 \times 2 \times 2$ kernel with a stride two operation along with the convolution layers in each stage of expansion paths.

5) MV_{NET}

The output of DA_{Net} is the input to MV_{Net} which has a similar 11 blocks as DA_{Net} along with a softmax layer as shown in Figure 3. Two feature maps are computed by the last $1 \times 1 \times 1$ convolution layer. The final output is the summation of two outputs from the convolution layer on par with the input image size. The classification of the pixels into MS lesions or normal tissue is done by applying a soft-max layer. The predicted output from MV_{Net} is,

$$\bar{I}_j^{(K-1)} = \text{Max}(0, \sum_{n=0}^{N-1} (\overline{W}_{n,i,j}^{(k)} * I_{(i)}^{(k-1)}) + a_{(j)}^{(k-1)}) \quad (7)$$

The output image from block 11 of MV_{Net} retains the same size as the input image.

IV. RESULTS AND DISCUSSION

All the models are implemented using Python version 3.9.13 and Keras 2.11 ran on intel(R) core (TM) i7 10870H

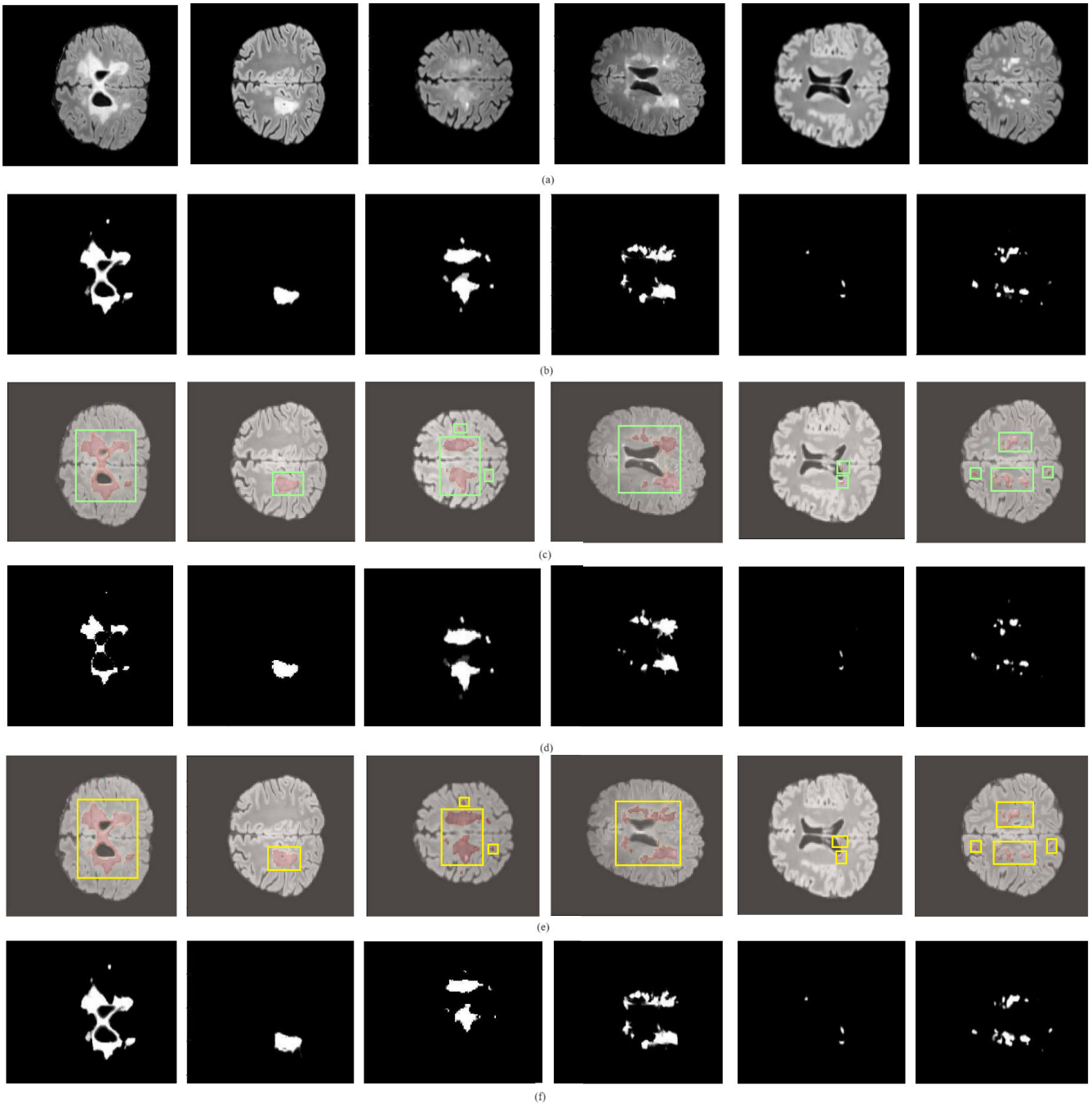

FIGURE 10. (a) Input MRI (b) Ground truth, (c) bounding box shows the diseased part using MV_{Net} (d) Output of MV_{Net}, (e) bounding box shows the diseased part using MV_{Net} (f) Output of DAV_{Net}.

CPU @ 2.20GHz 2.21 GHz and RAM is 16GB. The experiments were conducted on this model with 50 epochs, using an Adam optimizer with a learning rate of 0.001 and the loss function considered is the binary cross entropy loss. The stopping criteria is the early stopping criteria that is the model stops the the training process when the model performance on a validation set starts to worsen, preventing overfitting by selecting the model with the best generalization performance.

This section includes a dataset description, performance measures, and performance analysis. The following experiments were carried out to analyze the effectiveness of the DAV_{Net}. The first experiment evaluates the output of DA_{Net} (block 11) and the feature maps generated by block 12 of MV_{Net} in terms of heatmaps. The second experiment is the blockwise output analysis using entropy. The third experiment compares the performance of the DAV_{Net} with other DL models, the fourth experiment is the

TABLE 4. Entropy of output images of all the blocks.

Blocks of DA_{Net}	Entropy	Blocks of MV_{Net}	Entropy
1	2.31	12	3.49
2	2.96	13	3.31
3	3.30	14	3.42
4	3.39	15	3.93
5	4.01	16	4.77
6	4.12	17	4.60
7	5.22	18	5.27
8	4.73	19	4.77
9	3.69	20	4.14
10	3.48	21	3.68
11	2.77	22	0.77

comparison of DAV_{Net} with other MICCAI 16-based DL models, the fifth experiment is the comparison of DAV_{Net} with other DL models based on MICCAI 16 dataset, the sixth experiment is the performance analysis of DAV_{Net} based on MICCAI 16 and MSSEG -2 datasets, the seventh experiment is the comparison of DAV_{Net} with other DL models based on MSSEG-2 dataset and in the eighth experiment we have added the results of DAV_{Net} trained with MSSEG-2 dataset and tested on brain MRI MS-2022 dataset.

A. EXPERIMENT 1: ANALYSIS OF HEATMAPS OF BLOCK 11 AND BLOCK 12

This experiment is based on applying the proposed model to the MICCAI 16 dataset. As per Table 3, block 11 and 12 feature maps have the same resolution and block 12 delivers better feature maps than the feature maps from DA_{Net}. Figure 8 shows the heatmaps of different feature maps. Figure 8 (a) is the input, Figure 8 (b) is the heat map of block 11 feature maps, Figure 8 (c) is the heat map of block 12 input and Figure 8 (d) is the heat map of feature maps of block 12 outputs. by comparing Figure 8 (b) and (d), it is clear that the feature maps of block 12 better represent the details compared to block 11. The observation from this experiment is that more layers in a DL model can contribute more details to the segmentation process.

B. EXPERIMENT 2: BLOCK-WISE OUTPUT ANALYSIS USING ENTROPY

Figure 9 is an example of the output from different network layers of different blocks. The first row shows the output from blocks 1 to 11 of DA_{Net}, and the second row shows the output from blocks 12 to 22 (blocks 1 to 11 of MV_{Net}). Table 4 shows the corresponding image entropy of the successive blocks from two networks. While comparing the entropy of the grayscale images from different blocks of DA_{Net} with the corresponding block of MV_{Net}, it is observed that the entropy for the MV_{Net} blocks is higher. The MV_{Net} layers in different blocks have more details and contribute fine details for segmenting the region of interest. the output from the 22^d block is the segmented part since it is the binary image as shown in Figure 9(v) has low entropy.

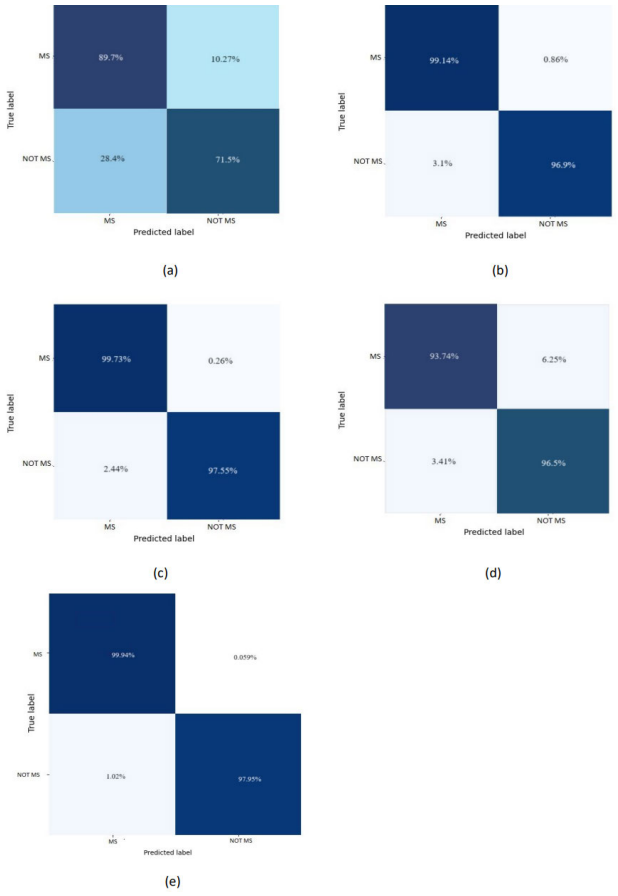

FIGURE 11. Confusion matrix corresponding to segmentation using (a) V-Net, (b) MV_{Net}: U Net, (c) U Net: MV_{Net}, (e) Sequential V-Net, and (e) DAV_{Net}.

TABLE 5. Comparison of the DAV_{Net} model with existing models.

DL model	Sensitivity	Precision	F1 Score	Loss	Mean IOU	Dice score
U-Net	0.920	0.954	0.503	0.070	0.250	0.533
ResNet	0.785	0.814	0.354	0.645	0.158	0.420
Inception V3-Net	0.416	0.496	0.333	0.694	0.110	0.350
DAV_{Net}	**0.9900**	**0.9930**	**0.8740**	**0.0007**	**0.8760**	**0.8900**

TABLE 6. Comparison of V-Net(N_1)Sequential V-Net (N_2), U-Net: MV_{Net} (N_3), MV_{Net}: U Net (N_4) and DAV_{Net} (N_5) models based on MICCAI 16 dataset.

Parameter	N_1	N_2	N_3	N_4	N_5
Sensitivity	0.7850	0.9518	0.9790	0.9850	0.9900
Precision	0.3970	0.9519	0.979	0.987	0.9930
Accuracy	0.7920	0.9518	0.980	0.987	0.9993
F1 Score	0.4063	0.4360	0.439	0.452	0.8740
Loss	0.4730	0.3983	0.467	0.385	0.0007
Mean IOU	0.1520	0.2605	0.250	0.324	0.8760
Dice sore	0.4960	0.5523	0.5589	0.5980	0.8900

C. EXPERIMENT 3: PERFORMANCE ANALYSIS DAV_{NET} WITH OTHER SEQUENTIAL MODELS

Figure 10 shows inputs and corresponding outputs of MV_{Net} and DAV_{Net}. MV_{Net} is a V-Net architecture with an additional

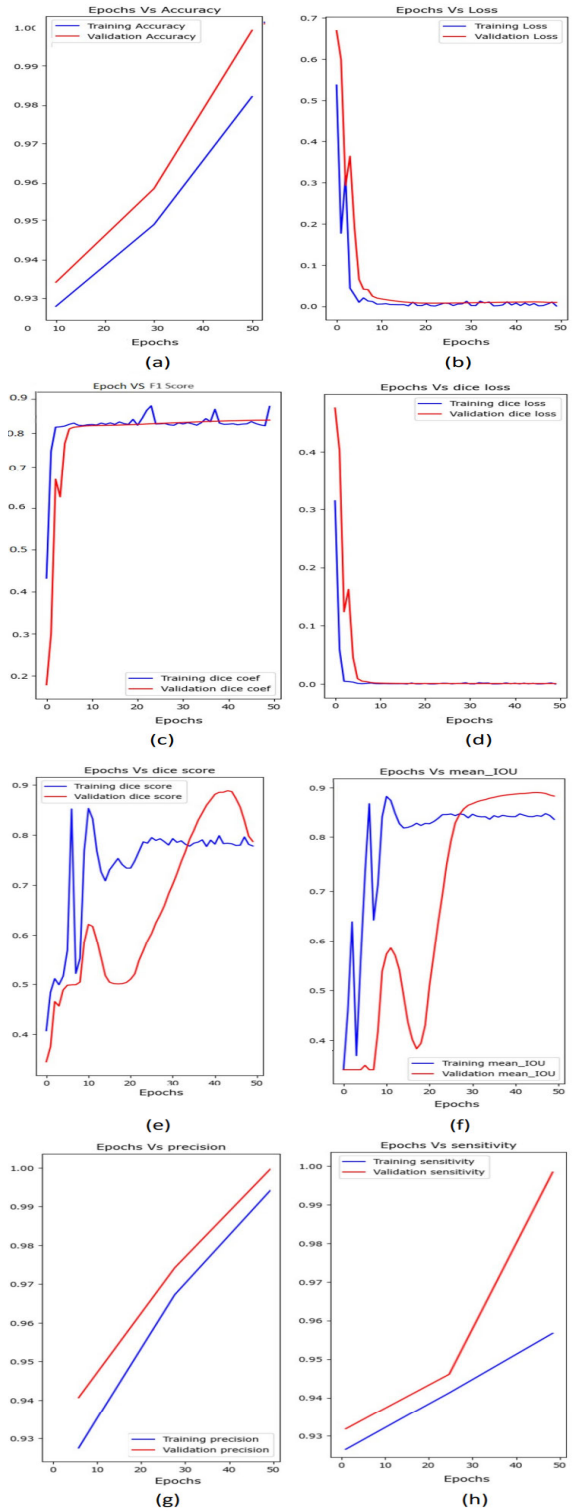

FIGURE 12. Performance measures of DAV_{Net}. (a) Epoch Vs accuracy (b)Epoch Vs loss (c) Epoch Vs F1 Score (d) Epoch Vs dice loss (e) Epoch Vs dice score (f) Epoch Vs mean IOU (g) Epoch Vs precision (h) Epoch Vs sensitivity.

block in the encoding and decoding paths. The segmented lesion appears closer to the ground truth lesion than the segmented MS lesion region from MV_{Net}. U-Net is a type of CNN used for medical image segmentation. Here, we have used U-Net with downsampling and upsampling layers for the segmentation. There are four downsampling blocks with series convolutional and max-pooling layers. convolutional layer extracts the feature map followed by ReLu activation and the maxpooling layer does the dimensionality reduction. Four upsampling blocks and one output block are present. The Up-sampling path has a concatenation layer, convolution, and upsampling layers. The output layer consists of the convolution with a softmax activation function. Residual Neural networks (ResNet) is also a DL architecture with a skip connection to learn residual mapping. Inception V3-Net is also a CNN architecture with encoding, decoding layers, skip connection and output layer. The performance measures of these DL models and the proposed DAV_{Net} model are presented in Table 5. From Table 5 it is clear that DAV_{Net} provides better segmentation results than the U-Net, ResNet, and Inception V3-Net models.

We have combined different DL models like U-Net and MV_{Net} in different ways to obtain the DL models like Sequential V-Net, U-Net: MV_{Net}, and MV_{Net}: U-Net. The above networks are tested for MS lesion segmentation in connection with the MICCAI 16 dataset. The V-Net has nine blocks and the MV_{Net} has an additional two blocks, blocks 6 and 7, as shown in Figure 3. This MV_{Net} serially connects with U-Net and generates two models such as MV_{Net}: U-Net and U-Net: MV_{Net}. Sequential V-Net is the DL model obtained by serially connecting two V-Net models having nine blocks. The analysis of the networks justifies the role of sequential arrangement of networks for improved segmentation.

1) CONFUSION MATRIX ANALYSIS

The confusion matrices for analyzing the performance of V-Net, MV_{Net}: U-Net, U-Net: MV_{Net}, sequential V-Net, and DAV_{Net} are presented in Figure 11. The confusion matrix presents the percentage of classified pixels as True Positive (TP), False Positive (FP), True Negative(TN), and False Negative (FN). The precision is the correctness of the positive prediction, By comparing the confusion matrices of the five models, V-Net has the least precision and a high number of FP. The number of FP of the MV_{Net}: U-Net is less than the V-net and higher than the U-Net: MV_{Net}. The precision of the MV_{Net}: U-Net is higher than the V-Net and less than the U-Net: MV_{Net}. The FP of sequential V-Net is higher than the MV_{Net}: U-Net. It is inferred from Table 6 and Figure 11 that as the FP increases, the precision of the DL model decreases.

D. EXPERIMENT 4: COMPARISON OF DAV_{NET} WITH OTHER DL MODELS

The following performance metrics are analyzed to compare predicted images with the ground truths.

$$Sensitivity(Recall) = TP/(TP + FN) \tag{8}$$

$$Precision = TP/(TP + FP) \tag{9}$$

$$Accuracy = TP + TN/(TP + TN + FP + FN) \tag{10}$$

TABLE 7. Results of MICCAI16 based DL models.

Reference	Details of network	Evaluation measures
Beaumont et al. [29]	A fully automated model with parameter optimization and lesion detection by graph cut initialized with a robust Expectation-Maximization (EM) algorithm. It removes FP.	Dice scores=0.5709, Sensitivity=0.3584, F1 Score= 0.3889
Beaumont et al. [30]	This method is based on image registration and intensity normalization. The refinement of the lesion after segmentation is done by lesion appearance rules.	Dice scores=0.5053, Sensitivity=0.5309, F1 Score = 0.2552
Knight et al. [31]	This is an edge-based model with partial averaging to estimate fuzzy membership profiles of tissue classes	F1 Score=0.60
Mahbod et al. [32]	It is a supervised artificial neural network with uses intensity-based features for lesion segmentation.	Dice score=0.66
Proposed model	Deep-attention V-Net architecture with modified compression and expansion sections	Dice score=0.890, Sensitivity=0.990, F1 Score=0.874

where the TP is the number of MS pixels correctly classified as MS, FP is the number of background pixels misclassified as MS, TN is the number of background pixels correctly classified as background and FN is the number of MS pixels misclassified as background. The Other two metrics in medical image segmentation are IOU (Intersection-over-Union) and F1 Score. The F1 Score is the harmonic mean between sensitivity and precision [1]. IOU penalizes under and over-segmentation more than the F1 Score. F1 Score is the most used metric in medical image segmentation.

$$IOU = TP/(TP + FP + FN) \tag{11}$$

$$F1\ Score = 2TP/(2TP + FP + FN) \tag{12}$$

The loss function considered here is the Binary cross-entropy loss. The binary cross entropy loss in the medical image segmentation compares the pixel-wise probabilities of input data and ground truth. It penalizes the model according to the deviation from the ground truth. DAV_{Net} delivered a precision of 0.9930, higher than all three models. The proposed method has a Dice score of 0.8900 and an F1 Score of 0.8740, higher than the existing three models. The proposed model achieves the lowest loss value compared to the three models. The sensitivity, specificity, and accuracy are higher for the proposed model than the existing DL models.

E. EXPERIMENT 5: COMPARISON OF DAV_{NET} WITH OTHER DL MODELS BASED ON MICCAI 16 DATASET

The models that were evaluated using dice score, F1 Score, and sensitivity are considered for comparative analysis and included in Table 7. Regarding dice score, sensitivity, and F1 Score, DAV_{Net} model-based segmentation delivers better results of 0.890,0.990, and 0.874 respectively.

F. EXPERIMENT 6: PERFORMANCE ANALYSIS DAV_{NET} BASED ON MICCAI 16 AND MSSEG-2 DATASETS

The DAV_{Net} model is evaluated on the combined MICCAI 16 and MSSEG-2. MICCAI 16 contains 53 NIfTI data and MSSEG-2 contains 100 data. So the dataset has a total number of 153 data. In this experiment, we used $n_{T1} = 70\%$ of data for training, $n_{T2} = 19\%$ of data for validation, and $n_{T3} = 11\%$ of data for testing. The performance metrics such as accuracy, loss, F1 Score, dice loss, mean IOU, dice score, precision, and sensitivity of the DAV_{Net} are evaluated and presented in Figure 12. It is observed that the number of epochs and loss function have opposite trends. As the epoch increases the accuracy also increases. Table 8 gives the evaluation result on test data.

TABLE 8. Test evaluation results corresponding to MICCAI 16 and MSSEG-2 datasets.

Parameter	Evaluation measures based on test data
Sensitivity	0.9997
Precision	0.9996
Accuracy	0.9997
Loss	0.0003
Mean IOU	0.8950
Dice sore	0.9100

G. EXPERIMENT 7: COMPARISON OF DAV_{NET} WITH OTHER DL MODELS BASED ON MSSEG-2 DATASET

MSSEG-2 dataset having 100 patients data obtained at two different time instants [33]. From the dataset. 40% data is used for training and 60% for testing. The values of the dice score compared with the results of the MSSEG-2 challenge are given in Table 9. Compared to the other models, the proposed model delivers a dice score of 0.9.

TABLE 9. Results of MSSEG-2 based DL models.

Reference	Dice score
Ashtari et al. [34]	0.63
Efird et al. [35]	0.54
La Rosa et al. [36]	0.36
Siddiquee et al. [37]	0.51
McKinley et al. [38]	0.56
Proposed model	**0.90**

H. EXPERIMENT 8: SEGMENTATION RESULTS OF DAV_{NET} WITH COMBINED MSSEG-2 AND MS MRI-2022 DATASETS

In this experiment, we have used two separate datasets. For training, we have used the MSSEG-2 dataset with 100 NIfTI data. The brain MRI MS 2022 dataset was used for validation and testing. 60 NIfTI data are available in the brain MRI MS 2022 dataset. Out of 60, 30 were used for validation and 30 for testing. The cross-validation results are presented in Table 10.

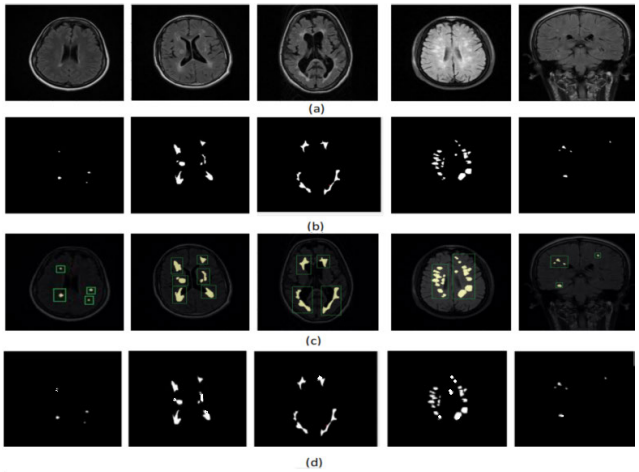

FIGURE 13. Segmentation results using combined MSSEG-2 and MRI MS 2022 dataset. (a) Input MRI,(b) Ground truth, (c) The bounding box shows the region of interest, (d) Outputs of DAV_{Net}.

This experiment is included in the study to generalize the model. Figure 13 shows a few output images and the performance measures corresponding to this experiment are presented in Table 10. Figure 14 shows the confusion matrix corresponding to experiment 8 regarding the number of pixels.

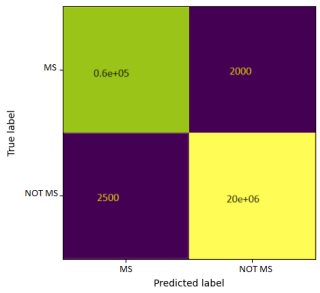

FIGURE 14. Confusion matrix corresponding to experiment 8.

TABLE 10. Performance measures corresponding to experiment 10.

Parameter	Validation
Accuracy	0.9997
Precision	0.9695
Sensitivity	0.9700
Loss	0.0003
Mean IOU	0.9300
Dice sore	0.9638

V. CONCLUSION

MS is a progressive illness of the CNS. It is due to demyelination and begins with the blood barrier breakdown. Even though remyelination occurs in the initial stage of the disease, nerve damage and irreversible loss of neurons occur in the initial phase of the disease. In this paper, a DL network, DAV_{Net}, is proposed for MS segmentation and tested on the MICCAI MRI data set. The experiments were conducted to analyze the role of the deep attention network and the effectiveness of the suggested model. The precision, sensitivity, accuracy, and loss based on the MICCAI 16 dataset are 0.9930, 0.9900, and 0.9993, 0.0007 respectively. The various experiments conducted using different datasets prove that the DAV_{Net} model performs well compared to the other DL models.

REFERENCES

[1] J. Hodler, R. A. Kubik-Huch, and G. K. von Schulthess, "Demyelinating diseases of the CNS (brain and spine)," in *Diseases of the Brain, Head and Neck, Spine 2020–2023*, vol. 1. Switzerland: Springer, 2020, pp. 165–170.

[2] A. Shoeibi, M. Khodatars, M. Jafari, P. Moridian, M. Rezaei, R. Alizadehsani, F. Khozeimeh, J. M. Gorriz, J. Heras, M. Panahiazar, S. Nahavandi, and U. R. Acharya, "Applications of deep learning techniques for automated multiple sclerosis detection using magnetic resonance imaging: A review," *Comput. Biol. Med.*, vol. 136, Sep. 2021, Art. no. 104697.

[3] T. Yousaf, G. Dervenoulas, and M. Politis, "Advances in MRI methodology," *Int. Rev. Neurobiol.*, vol. 141, pp. 31–76, Jan. 2018.

[4] S. S. Kunapuli and P. C. Bhallamudi, "A review of deep learning models for medical diagnosis," in *Machine Learning, Big Data, and IoT for Medical Informatics*, 1st ed., Solan, India: Jaypee University of Information Technology, Academic, 2021, pp. 389–404.

[5] A. Mathew, P. Amudha, and S. Sivakumari, "Deep learning techniques: An overview," in *Proc. Adv. Mach. Learn. Technol. Appl. (AMLTA)*, 2021, pp. 599–608.

[6] N. Siddique, S. Paheding, C. P. Elkin, and V. Devabhaktuni, "U-net and its variants for medical image segmentation: A review of theory and applications," *IEEE Access*, vol. 9, pp. 82031–82057, 2021.

[7] H. Jiang, Z. Diao, T. Shi, Y. Zhou, F. Wang, W. Hu, X. Zhu, S. Luo, G. Tong, and Y.-D. Yao, "A review of deep learning-based multiple-lesion recognition from medical images: Classification, detection and segmentation," *Comput. Biol. Med.*, vol. 157, May 2023, Art. no. 106726.

[8] N. Dong, L. Zhao, C. H. Wu, and J. F. Chang, "Inception V3 based cervical cell classification combined with artificially extracted features," *Appl. Soft Comput.*, vol. 93, Aug. 2020, Art. no. 106311.

[9] Z. Lu, Y. Bai, Y. Chen, C. Su, S. Lu, T. Zhan, X. Hong, and S. Wang, "The classification of gliomas based on a pyramid dilated convolution resnet model," *Pattern Recognit. Lett.*, vol. 133, pp. 173–179, May 2020.

[10] C.-Y. Lee, S. Xie, P. Gallagher, Z. Zhang, and Z. Tu, "Deeply-supervised nets," in *Proc. 18th Int. Conf. Artif. Intell. Statist.*, 2015, pp. 562–570.

[11] A. Carass et al., "Longitudinal multiple sclerosis lesion segmentation: Resource and challenge," *NeuroImage*, vol. 148, pp. 77–102, Mar. 2017.

[12] E. G. Kholmovski, A. A. Samsonov, and D. L. Parker, "Motion artifact reduction technique for dual-contrast FSE imaging," *Magn. Reson. Imag.*, vol. 20, no. 6, pp. 455–462, Jul. 2002.

[13] S. Vaidya, A. Chunduru, R. Muthuganapathy, and G. Krishnamurthi, "Longitudinal multiple sclerosis lesion segmentation using 3D convolutional neural networks," in *Proc. Longitudinal Multiple Sclerosis Lesion Segmentation Challenge*, 2015, pp. 1–2.

[14] S. Valverde, M. Salem, M. Cabezas, D. Pareto, J. C. Vilanova, L. Ramió-Torrentà, À. Rovira, J. Salvi, A. Oliver, and X. Lladó, "One-shot domain adaptation in multiple sclerosis lesion segmentation using convolutional neural networks," *NeuroImage: Clin.*, vol. 21, 2019, Art. no. 101638.

[15] F. Milletari, S.-A. Ahmadi, C. Kroll, A. Plate, V. Rozanski, J. Maiostre, J. Levin, O. Dietrich, B. Ertl-Wagner, K. Bötzel, and N. Navab, "Hough-CNN: Deep learning for segmentation of deep brain regions in MRI and ultrasound," *Comput. Vis. Image Understand.*, vol. 164, pp. 92–102, Nov. 2017.

[16] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "NnU-net: A self-configuring method for deep learning-based biomedical image segmentation," *Nature Methods*, vol. 18, no. 2, pp. 203–211, Feb. 2021.

[17] T. Brosch, L. Y. W. Tang, Y. Yoo, D. K. B. Li, A. Traboulsee, and R. Tam, "Deep 3D convolutional encoder networks with shortcuts for multiscale feature integration applied to multiple sclerosis lesion segmentation," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1229–1239, May 2016.

[18] C. Zhao, B. E. Dewey, D. L. Pham, P. A. Calabresi, D. S. Reich, and J. L. Prince, "SMORE: A self-supervised anti-aliasing and super-resolution algorithm for MRI using deep learning," *IEEE Trans. Med. Imag.*, vol. 40, no. 3, pp. 805–817, Mar. 2021.

[19] G. Kang, B. Hou, Y. Ma, F. Labeau, and Z. Su, "Acu-net: A 3D attention context U-net for multiple sclerosis lesion segmentation," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 1384–1388.

[20] E. Roura, A. Oliver, M. Cabezas, S. Valverde, D. Pareto, J. C. Vilanova, L. Ramió-Torrentà, À. Rovira, and X. Lladó, "A toolbox for multiple sclerosis lesion segmentation," *Neuroradiology*, vol. 57, no. 10, pp. 1031–1043, Oct. 2015.

[21] P. Schmidt, C. Gaser, M. Arsic, D. Buck, A. Förschler, A. Berthele, M. Hoshi, R. Ilg, V. J. Schmid, C. Zimmer, B. Hemmer, and M. Mühlau, "An automated tool for detection of FLAIR-hyperintense white-matter lesions in multiple sclerosis," *NeuroImage*, vol. 59, no. 4, pp. 3774–3783, Feb. 2012.

[22] O. Commowick, M. Kain, R. Casey, R. Ameli, J.-C. Ferré, A. Kerbrat, T. Tourdias, F. Cervenansky, S. Camarasu-Pop, T. Glatard, S. Vukusic, G. Edan, C. Barillot, M. Dojat, and F. Cotton, "Multiple sclerosis lesions segmentation from multiple experts: The MICCAI 2016 challenge dataset," *NeuroImage*, vol. 244, Dec. 2021, Art. no. 118589.

[23] R. McKinley, R. Wepfer, T. Gundersen, F. Wagner, A. Chan, R. Wiest, and M. Reyes, "Nabla-net: A deep dag-like convolutional architecture for biomedical image segmentation," in *Proc. Brainlesion: Glioma, Multiple Sclerosis, Stroke, Traumatic Brain Injuries, 2nd Int. Workshop, BrainLes, Challenges BRATS, ISLES mTOP, Held Conjunction (MICCAI)*, Athens, Greece, Springer, 2016, pp. 119–128.

[24] S. R. Hashemi, S. S. Mohseni Salehi, D. Erdogmus, S. P. Prabhu, S. K. Warfield, and A. Gholipour, "Asymmetric loss functions and deep densely-connected networks for highly-imbalanced medical image segmentation: Application to multiple sclerosis lesion detection," *IEEE Access*, vol. 7, pp. 1721–1735, 2019.

[25] P. Ghosal, P. K. C. Prasad, and D. Nandi, "A light weighted deep learning framework for multiple sclerosis lesion segmentation," in *Proc. 5th Int. Conf. Image Inf. Process. (ICIIP)*, Nov. 2019, pp. 526–531.

[26] H. M. Afzal, S. Luo, S. Ramadan, J. Lechner-Scott, M. R. Amin, J. Li, and M. K. Afzal, "Automatic and robust segmentation of multiple sclerosis lesions with convolutional neural networks," *Comput., Mater. Continua*, vol. 66, no. 1, pp. 1–15, 2021.

[27] R. McKinley, R. Wepfer, F. Aschwanden, L. Grunder, R. Muri, C. Rummel, R. Verma, C. Weisstanner, M. Reyes, A. Salmen, A. Chan, F. Wagner, and R. Wiest, "Simultaneous lesion and brain segmentation in multiple sclerosis using deep neural networks," *Sci. Rep.*, vol. 11, no. 1, p. 1087, 2021.

[28] G. Latif, J. Alghazo, F. N. Sibai, D. N. F. A. Iskandar, and A. H. Khan, "Recent advancements in fuzzy C-means based techniques for brain MRI segmentation," *Current Med. Imag.*, vol. 17, no. 8, pp. 917–930, 2021, doi: 10.2174/1573405616666210104111218.

[29] J. Beaumont, O. Commowick, and C. Barillot, "Multiple Sclerosis lesion segmentation using an automated multimodal graph cut," in *Proc. 1st MICCAI Challenge Multiple Sclerosis Lesions Segmentation Challenge Using Data Manag. Process. Infrastruct. (MICCAI-MSSEG)*, Oct. 2016, pp. 1–8.

[30] J. Beaumont, O. Commowick, and C. Barillot, "Automatic multiple Sclerosis lesion segmentation from intensity-normalized multi-channel MRI," in *Proc. 1st MICCAI Challenge Multiple Sclerosis Lesions Segmentation Challenge Using Data Manag. Process. Infrastruct. (MICCAI-MSSEG)*, 2016.

[31] J. Knight and A. Khademi, "MS lesion segmentation using FLAIR MRI only," in *Proc. 1st MICCAI Challenge Multiple Sclrosis Lesions Segmentation Challenge Using Data Manag. Process. Infrastruct. (MICCAI-MSSEG)*, 2016, pp. 21–28.

[32] A. Mahbod, C. Wang, and O. Smedby, "Automatic multiple sclerosis lesion segmentation using hybrid artificial neural networks," in *Proc. 1st MICCAI Challenge Multiple Sclerosis Lesions Segmentation Challenge Using Data Manag. Process. Infrastruct. (MICCAIMSSEG)*, 2016, pp. 29–36.

[33] O. Commowick, F. Cervenansky, and R. Ameli, "MSSEG challenge proceedings: Multiple sclerosis lesions segmentation challenge using a data management and processing infrastructure," in *Proc. 24th Int. Conf. Med. Image Comput. Comput. Assist. Intervent. (MICCAI)*, 2016, p. 126.

[34] P. Ashtari, B. Barile, S. Van Huffel, and D. Sappey-Marinier, "Longitudinal multiple sclerosis lesion segmentation using pre-activation U-Net," in *Proc. Challenge: Muliple Sclerosis New Lesions Segmentation Challenge Using Data Manag. Process. Infrastruct. (MSSEG)*, 2021, pp. 45–51.

[35] C. Efird, D. Miller, and D. Cobzas, "A UNet pipeline for segmentation of new MS lesions," in *Proc. Challenge: Multiple Sclerosis New Lesions Segmentation Challenge Using Data management Process. Infrastruct. (MSSEG)*, 2021, p. 53.

[36] F. La Rosa, J. P. Thiran, and M. Bach, "A subtraction image-based method to detect new appearing multiple sclerosis lesions on single-contrast FLAIR MRI," in *Proc. Challenge: Multiple Sclerosis New Lesions Segmentation Challenge Using Data Manag. Process. Infrastruct. (MSSEG)*, 2021, p. 65.

[37] M. M. R. Siddiquee and A. Myronenko, "Robust 3D MRI segmentation of multiple sclerosis lesions," in *Proc. Challenge: Multiple Sclerosis New Lesions Segmentation Challenge Using Data Manag. Process. Infrastruct. (MSSEG)*, 2021, p. 81.

[38] R. McKinley, F. Wagner, and R. Wiest, "Detection of lesion change in multiple sclerosis using a cascade of 3D-to-2D networks," in *Proc. Challenge: Multiple Sclerosis New Lesions Segmentation Challenge Using Data Manag. Process. Infrastruct. (MSSEG)*, 2021, p. 97.

V. P. NASHEEDA received the B.Tech. degree in electronics and communication engineering and the M.Tech. degree in communication engineering and signal processing from the University of Calicut, Kerala, India, in 2011 and 2014, respectively. She is currently pursuing the Ph.D. degree in medical image processing with the School of Electronics, Vellore Institute of Technology, Chennai, India. From 2015 to 2019, she was an Assistant Professor with the Electronics and Communication Engineering Department, Kerala Technical University, India, in various engineering colleges.

VIJAYARAJAN RAJANGAM (Senior Member, IEEE) received the Bachelor of Engineering degree in electronics and communication engineering from the University of Madras, in 1998, the master's degree in applied electronics from Madurai Kamarajar University, in 1999, and the Ph.D. degree in image fusion from Anna University, in 2015. He is currently an Associate Professor with the Centre for Healthcare Advancement, Innovation, and Research, School of Electronics Engineering, VIT, Chennai. He has been in academics for more than 23 years and has published 42 research contributions in various SCI/Scopus journals and conferences. He is guiding Ph.D. scholars in the field of medical image analysis. His research interests include computer vision, machine learning algorithms for signal and image processing, image fusion, bio-cryptography, emotion classification, and deep learning for image analysis. He is a fellow of IETE and a Life Member of ISTE.