MJCIS Vol.15 No.2 Dec 2019

Mansoura Journal of Computers and Information Sciences

Automatic Multiple Sclerosis lesion segmentation using Patch-wise R-CNN

Doaa Al-Desouky
Faculty of Computers and Information,
Computer Science Dept.
Mansoura University, Egypt
doaa.aldesouky@gmail.com

Ehab Essa
Faculty of Computers and Information,
Computer Science Dept.
Mansoura University, Egypt
ehab_essa@mans.edu.eg

M. Z. Rashad
Faculty of Computers and Information, Computer Science Dept.
Mansoura University, Egypt
magdi_z2011@yahoo.com

Sherif E. Hussein
Faculty of Engineering, Computer and Systems Dept.
Mansoura University, Egypt
sherif_hussein@mans.edu.eg

ABSTRACT

Multiple sclerosis (MS) could be considered one of the most severe neurological diseases, which can cause damage to the central nervous system. Because of the regular change in size, location and anatomical variation of MS lesions, it is a challenge to accurately identify, characterize and quantify MS lesions on magnetic resonance imaging (MRI). Therefore, MS lesion segmentation and detection become an active point of research.

Recently, deep neural networks (DNN) have seen a rapid advance in various medical image analysis fields, i.e., image registration, image segmentation, lesion detection, and shape modeling. Furthermore, convolution neural networks (CNN) have gained popularity in medical imaging, especially in brain imaging.

In this study, an automated technique is proposed to segment MS lesions in MRI. This technique depends on a 3D patch-wise region-based convolution neural network (R-CNN) for MS lesion segmentation in T2-w and FLAIR.

The proposed method is evaluated using the public MICCAI2008 MS lesion segmentation data set, which is compared to other MS lesion segmentation tools.

General Terms
Deep learning

Keywords
R-CNN; Deep learning; MS Segmentation.

1. INTRODUCTION

Multiple sclerosis (MS) is a chronic disease of the central nervous system (CNS), which causes the immune system to attack the protective sheaths of myelin around the axons in the brain. MS infects over 2.5 million individuals, as it is the primary cause of non-traumatic neurological paralysis in youth in Europe and North America, [1]. The pathological symbol of MS is the focal area's existence of inflammatory-mediated spinal cord white matter (WM) and demyelination of the brain. Therefore, symptoms of MS differ from a patient to another according to the size, number, and location of the affected lesion. Some patients have moderate symptoms which do not need medication, while, others may have a physical disability, cognitive decline problems, problems in moving around and doing the daily activities. The reason for MS is unknown until now; many researchers proposed that the disease could result from the complex interaction between environmental, genetic and immunological issues.

In the early 1980s, MRI was presented to evaluate and diagnose MS [2]. MRI has become the most valuable mechanism for evaluating MS patients, as it not only helps in diagnosis but also monitors the progress of the disease and determines the pattern of the disease. Conventional MRI such as proton density-weighted (PD), fluid-attenuated inversion recovery (FLAIR), T1-weighted (T1) and T2-weighted (T2) are sensible to lesions and demonstrate these lesions with various intensity to peripheral tissues. Further, the intensity of MS lesions and the brain tissue are various in different MRI sequences; therefore, to accomplish an adequate algorithm, it is essential to use multispectral images. Segmentation techniques depend primarily on the quality, application, and characteristics of MRI. Segmentation of MS lesions in MRI is challenging because of various reasons, for instance, irregular intensity, partial volume effects, noise and the parameters of imaging [3].

Although manual segmentation of MS lesions is time-consuming besides it focuses on inter- and intra –expert variability, it is still acknowledged as the MS standard. Therefore, over the last 20 years, the necessity to develop a fully automated MS lesion segmentation technique becomes vital in the medical imaging community research [4]. The validation of segmentation in medical images is complicated because of the deficiency of an adequate standard reference required to compare the results of any segmentation approach. It is helpful to compare the segmentation results to histology, however, for clinical data, it is scarcely available, and it can be challenging to relate histology to MRI [5]. Hence, in general, validation studies depend on expert imaging data evaluation.

2. RELATED WORKS

Nowadays, standard techniques used for image analysis in medical trials are manual; which is expensive, complicated and time-consuming. In manual segmentation, errors exist because of partial volume effects; indefinite borders resulted from the change in tissue characteristics and low lesion contrast. Inconsistencies in the segmentation are also common, including experienced specialists. Several researchers have studied the variability ingrained to manual segmentation of MS lesion who stated that the intra-rater volume variability of 6.5% and inter-rater of 14% [6]. Hence, an automatic multitude of methods for MS lesion segmentation and detection have been suggested [7]. The recommended methods for segmentation of MS lesions comprise supervised and unsupervised learning.

Unsupervised methods are suitable for practical employment since they do not participate in previous knowledge. Unsupervised lesion segmentation methods are based on brain tissue intensity models, where the image voxels, including high intensities in FLAIR, are modeled as abnormal values that depend on intensity distributions [8]. Optimal segmentation finally has achieved via an expectation maximization algorithm. After that, the outlier voxels turned into the potential candidates for lesions and later a simple threshold can improve the segmentation [9]. Instead, Bayesian systems such as mixtures of Gaussians [10] or Student's mixture models [11] characterized lesions as outliers of the mixture model. In addition, Garc´ıa-Lorenzo et al. [12] joint this technique with a mean shift algorithm to segment the MS lesion.

Further information on the intensity distribution and the predicted location of normal tissues can be included through a collection of healthy subjects [2] to determine lesions with greater accuracy. Local intensity information may also be included through a Markov random field (MRF) to achieve soft segmentation [13]. The detected lesions in the range of the white matter (WM) using the mathematical morphology method of coarse-to-fine of T1-w, FLAIR, and diffusion-weighted imaging (DWI) sequences. In [14], fuzzy clustering specifies WM voxels and outlier gray matter (GM), and after that, lesion maps are being produced. In order to correct for noise and image artifacts, graph-cut techniques are applied for combining spatial information from the local neighborhoods with the intensity model [12]. One of the principal challenges of unsupervised methods is that, frequently, the outliers are not specific to lesions, which may be due to structures of a small anatomical, for instance, blood vessels, imaging artifacts, partial volume, and intensity inhomogeneity which lead to false positive results.

Supervised methods for lesion segmentation apply atlases or templates, which in general consist of multi-contrast MRI besides their manually delineated lesions. As seen in the ISBI-2015 lesion segmentation challenge [15], supervised techniques have become more common and significant than unsupervised methods, with 4 of the top 5 supervised methods. Supervised techniques learn how to transform the intensity of MRI into lesion labels in atlases. After that, the learned transformation is implemented to a new invisible image to generate the lesion labels of this image. Logistic regression [16] and support vector machines [17] was applied for lesion classification, where features include voxel intensity of the voxel-wise from multi-contrast images and the task of classification is to label the voxel image as a lesion or not. In [18], the authors used k-NN classification with tissue type before classifying WM lesions from 3.0T MRI. In [19], MS lesions in FLAIR images with noise, also, other types of artifacts are automatically segmented using the Cellular Neural Network. Another method proposed in [20] requires the user to supply some ROI's (Regions of Interest), which include lesions as the ground truth rather than the labeled lesion map. As an alternative to the voxel-wise intensities, patches have proven to be a convenient and robust feature. Roy et al. [21] proposed patch-based lesion segmentation using samples from an atlas corresponding to patches in input images by applying a sparse dictionary method. The algorithms based on random forests and k-NN [22, 23] utilized patches besides other features, calculated at a specific voxel, for predicting the label of that voxel. The dictionary-based approaches [24-26] apply patches of an image from atlases to learn a patch dictionary, which adequately defines patches of potential lesions and non-lesions. For a new unseen patch, identical patches are discovered from the dictionary and associated with weights based on the similarity.

The recent advance in automatic segmentation applying deep learning originates from a domain of the cell membrane segmentation proposed by Cires¸an et al. [27]. They had suggested applying a convolution neural network (CNN) to classify the image patch centers without the feature extraction step. Alternatively, the lower layer of the network learns the features indirectly during the training, whereas the higher layers perform the classification using the learning ability of features. Despite, if the number and size of patches are large, the time required for training makes the approach not applicable.

Recently, CNN or deep learning [28], have achieved sophisticated results in various computer vision challenges like object recognition and detection. CNN become popular in brain imaging, particularly in tissue segmentation [29] and brain tumor segmentation [30, 31]. The principal advantage of CNN over traditional machine learning algorithms is that CNN does not require hand-made features, so it is suitable for a different set of problems when finding the optimal features is unclear. CNN can manage 3D images or image patches. Therefore, 2D [35] and 3D [34] algorithms have been proposed, but 2D patches are often favored because of the speed and memory efficiency. With progress in graphics processor units (GPU), neural network models can be trained on a GPU within a fraction of time taken by that with multiple CPUs. Brosch et al. [34] submitted a cross-sectional method to segment MS lesions rely on deep 3D CNN with two interconnected pathways and shortcut connections. Havaei et al. [31] proposed a segmentation structure for lesions including convolution pipes of autonomous image modality to diminish the missing modalities impact of new unseen instances. Maleki et al. [36] used the CNN architecture to extract features from 2D MR FLAIR images and a three-layer neural network as a classifier.

This paper proposes a patch-wise region-based convolution network (R-CNN) for the automated segmentation of MS lesion. First, the input images are pre-processed to reduce noisy artifacts and as a result, enhance the segmentation accuracy. For this purpose, the skull stripping, and bias field correction have been applied. After that, 3D patches have been extracted from the input images for training the R-CNN model and producing a probability map of the MS lesion. The proposed method has been tested on the publicly available MICCAI2008 MS challenge dataset to differentiate its performance from other MS lesion segmentation methods.

3. THE PROPOSED METHOD

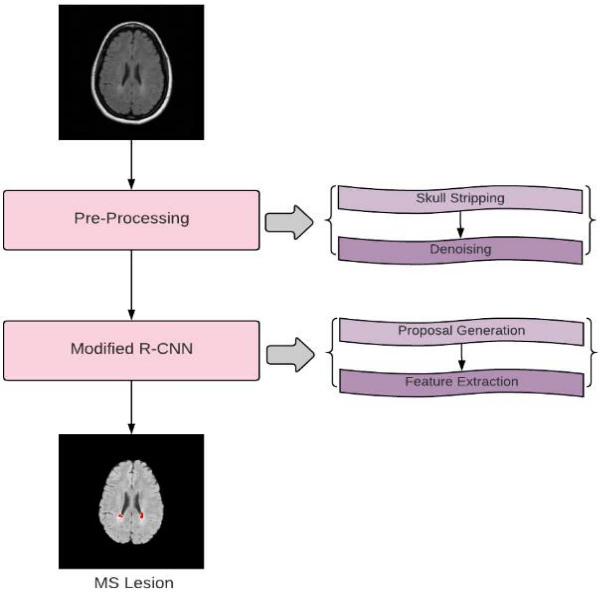
Since 2012, CNN has gained much acceptance after the ImageNet classification challenge, where Alex Krizhevsky's network obtained a wide margin [37]. Since then, CNN has been used favorably for additional applications like object segmentation and detection. Because of the enormous number of CNN parameters and the compact structure, recognizing the reason why they perform so well is not identified, numerous researches have been committed for this purpose. Recently, CNN has been utilized favorably in an analysis of medical images, wherever volumetric data are usually obtainable.

CNN learns the relations between the pixels of input images by obtaining the representative features through the convolution and pooling processes. The features detected in every layer that use the learned kernels differ in terms of complexity, that is to say, the first layers extract the simple features, such as the edges and the successive layers, extracting more complex and high-level features. The convolution procedure in CNN has three essential benefits. First of all, the weight sharing technique allows managing high dimensional data, whether 2D or 3D, like videos and volumetric images. Secondly, the local connectivity of input topology can be employed using 2D or 3D kernels. Finally, a slight change invariance is obtained by applying the pooling layer.

The proposed system overview is shown in Figure 1. First, the input of modality FLAIR or T2 is pre-processed. Second, a patch-wise technique used for region proposal generation and finally, extract a fixed-length feature vector from every region and fed it to the R-CNN.

Figure 1. The proposed system overview

The initial step aims to pre-process the original images. In the pre-processing phase, the skull stripping was performed in addition to the bias field correction. After pre-processing the



images, 3D patches were obtained from the input images obtained for training an R-CNN model and producing a probability map of the MS lesion. The proposed method has been tested using MICCAI 2008; the MS challenge data set compares its performance and other MS lesion segmentation methods.

3.1 Pre-processing

Segmentation of the brain MRI is complicated due to various reasons, such as normal anatomical variations, blurred edges, noise, overlapping intensities, and the variable imaging parameters [8]. Thus, pre-processing steps should be performed before implementing any strategy for MS lesion segmentation. Pre-processing of MRI has a significant influence on the accuracy of segmentation results, which composed of these steps: co-registration, skull stripping, bias field correction, normalization of the intensity as well as noise reduction. It is significant to perform these steps respectively to improve the influence of image quality. The MICCAI 2008 dataset have been previously co-registered, which refers that the same voxel in various sequences indicates an identical location in the brain.

3.1.1 Skull stripping

Magnetic resonance volumes in the brain incorporate parts of the non-brain tissue of the head, e.g., the spinal cord, eyes or skull. Removing the brain tissue from a non-brain tissue called skull stripping. For this purpose, an automated brain extraction algorithm was applied to images using brain extraction tools (BET) in the MRIcro software [38].

3.1.2 Denoising and bias-field correction

The inherent characteristics of the MRI acquisition process, such as eddy currents driven by field gradients, bandwidth filtering of data, or differences in the magnetic field lead to image artifacts, which itself leads to a negative influence in the performance of methods [44 39]. Spurious intensity variations should be eliminated due to lack of homogeneity of magnetic fields and coils. At such states, a correction of the magnetic resonance intensity is implemented as part of the tissue segmentation pipeline or before tissue segmentation. The well-known procedure for solving that issue is applying the field correction for a bias. Image inhomogeneities occur when the same biological tissues created by the bias field have a different intensity. In this case, the N3 method proposed in [40] to estimate and correct these inhomogeneities. Moreover, the filter of anisotropic diffusion was used to lower the noise without blurring insignificant morphological characteristics [41]. To normalize the intensity of the image profiles (from 0 to 255), the docile based piecewise linear transformation method is used. Magnetic resonance data sets of the brain can have volumes obtained from different scanner manufacturers or even from the same scanner but with different protocols. Hence, volumes can show a representation of the non-uniform intensity of the same models of tissue, that is, the inter-class variability.

3.2 Convolution Neural Network (CNN) architecture

CNN was introduced in 1989 [42], but in 2012 it gained an excessive interest, especially when deep CNN had achieved tremendous results in the Image Net competition [37]. CNN has been used in many data sets of millions of images, including many different classes. CNN has almost decreased the error rates of the previous top computing methods [37]. The architecture of CNN is usually complex since in some structures it can have more than a hundred layers, millions of weights, and billions of neuron connections. The architecture usually includes convolution layers, pooling layers, activation layers, and classification layers. The convolution layer generates feature maps by involving a kernel through an input image. The pooling layer has applied for reducing the output of previous convolution layers. To this end, an average or maximum of a specified neighborhood is applied as a transferred value to the following layer. The Rectified Linear Unit (ReLU) beside its modifications, like Leaky ReLU, is considered as activation functions. ReLU converts data non-linearly by decreasing negative input values to zero, but the positive is transferred as an output value [43].

The output of the last CNN layer is linked with the loss function (such as cross-entropy) to predict the input data. Finally, the network parameters are identified by decreasing the loss function among the ground truth and prediction labels with regularization constraints, whereas, network weights are updated at every iteration by applying the back propagation till convergence.

3.3 Modified R-CNN

Many procedures have been indicated in computer vision research for region proposals. R-CNN [44] outperforms traditional methods in PASCAL VOC, benefiting from these two ideas: First, object proposals have been applied instead of sliding windows. R-CNN has proposed a specified number of boxes for all images that probably has target objects. The various scales' problem is managed automatically by generating the proposal. The lesser proposals, although better, shows offer more for the superior performance of R-CNN. Secondly, the pre-trained DNN prototypes of Image Net were used, which were subsequently fine-tuned by applying PASCAL VOC. The procedure of pre-training is approved as essential for performance. Given proposals of the region, training of the R-CNN object detector includes two main stages: domain-specific fine-tuning and supervised pre-training.

Fast R-CNN [45] has adjusted some of the weaknesses of R-CNN, by proposing a layer of ROI Pooling, which produces proposals from the feature map rather than from the input image. It also presents a regression step for refining the region's box. Faster R-CNN [46] presented a common end-to-end training for a region's classification module. Moreover, the CNN region proposal module (applied a selective search for region proposal instead) included a substantial weight sharing, which improved the quality and speed of the detection than the original R-CNN [44].

Despite that, Fast R-CNN reduces the cost of time as well as enhances the performance of the VOC PASCAL; the central concept of R-CNN is robust. The appending of ROIPooling results in the main distinction between these two approaches. In the R-CNN, all proposal boxes (including the small) have been resized to an approved dimension what indicates that a full map of features is created for every proposal box in the last pooling layer. While, in the Fast R-CNN, a small proposal box is assigned to a small map in the final pooling layer. This feature map might scarcity the essential data for the classification phase, which adds doubts to the research. Therefore, the R-CNN is adequate to our proposed method rather than the Fast R-CNN procedure. For example, visualizing the neuronal responses of R-CNN is further appropriate than the Fast R-CNN. Moreover, interpreting the impact of the up-sampling or context is simpler when working with the proposal patch input. Therefore, for MS lesion segmentation in the proposed method, the original R-CNN pipeline has been chosen. Despite being a regression of the bounding box is a more efficient method for enhancing the accuracy of localization, it is not the main difficulty in lesion segmentation; hence the bounding box regression has not implemented.

In the proposed method, R-CNN was adopted to segment MS lesion. The 3D patches were extracted from the input MRI modalities to work as region proposal. The R-CNN was modified by adding a softmax layer to classify the input patch as a lesion or not. The output was used to generate a probability map of the lesion locations; as a result, there was no need for either a boundary box regression model or post-processing classification. The proposed R-CNN method follows three phases; initially, region proposals were generated to localize MS lesions. Then, R-CNN architecture was built. Also, both the transfer learning and fine-tuning were applied to train the model. Finally, the probability map of lesion locations was generated.

4. TRAINING DETAILS

In this section, the details of the proposed lesion segmentation model are covered.

4.1 Proposal Generation

Although there are several methods to generate region proposals, for example, selective search [47], a patch-wise technique was used in the proposed method. The patch was considered from the specified MRI by placing a window of size 7×7×3 around each pixel, which then is normalized to unit standard deviation and zero mean. Each patch label was specified by the center pixel.

4.2 Feature extraction

The fixed-length feature vector was extracted using the CNN per region proposal. To compute the region proposal features, first, the image data in this region was converted into a compatible format with CNN. Therefore, an input patch was to 32 × 32 × 3 to take advantage of the deep R-CNN architecture.

4.3 Training procedure

First, CNN was pre-trained on a large set of labeled data; CIFAR-10, which composed 32x32 images related to 10 object classes. The proposed CNN model has one fully connected layer, three pooling layers, and three convolution layers. The first block of CNN is a convolution layer with a 5 x 5 kernel that generates 32 feature maps and a max-pooling layer with a 3 x 3 kernel. These three levels are utilized for detection of low-level features in the patch. Subsequent, these layers are trailed with another block of a convolution layer, RELU and a max-pooling layer including the same kernels, which generate 32 feature maps that utilized for the detection of higher-level image structures. Following that, another convolution layer was applied with a 5 x 5 kernel, which generates 64 feature maps. The remainder of CNN has included two fully connected layers whose aim is to use these features to classify the input image to several classes. To generate a probabilistic lesion output, a softmax loss layer of size 2 was added. The ReLU function was applied as an activation function in every convolution layer. In this experiment, the learning rate was specified at 0.0001, which then is multiplied by 0.2 for every 15 epochs passed during the training. Then, fine-tuning this trained CNN on MS lesion dataset. To adapt the CNN to the MS lesion segmentation and the patches of MRI, the stochastic gradient descent (SGD) proceeded to train the CNN parameters utilizing the patch region proposals. The region proposal was handled that overlaps with ≥ 0.6 with the ground-truth as positive while otherwise that, is as negative. To detect MS lesions, consider the training as a binary classifier, and a positive example is the image region tightly surrounding a lesion while the background is negative. In order to label a region, the IoU overlap threshold was applied, where below than 0.1 were defined as negative examples, whereas over than 0.6 could be defined as positive. The overlap threshold was named by a grid search in the validation set. It is essential to select the threshold carefully. After features extraction and applying training labels, the softmax loss layer of size 2 was optimized to generate a probabilistic output of lesion existence.

5. EXPERIMENTS AND RESULTS

5.1 Data

The MICCAI 2008 MS lesion segmentation challenge includes 45 scans from research subjects assimilated at University of North Carolina (UNC, 3T Siemens Alegra) and Children's Hospital Boston (CHB, 3T Siemens) ([32]). For every subject, the image patterns T1-w, T2-w and FLAIR are delivered with an isotropic resolution of 0.5 × 0.5 × 0.5 mm3 in all images. Two datasets:

- 20 training cases (10 from CHB whereas 10 from UNC) with the manual annotations of WM lesions by a CHB and UNC experts
- 25 test cases (15 CHB and 10 UNC) without expert segmentation of the lesion.

5.2 Evaluation

The evaluation was blindly completed for the teams by delivering the segmentation masks of 25 test cases to the challenge website. The provided segmentation masks were compared to the manual annotations between the CHB and UNC evaluators. The following scores are based on the evaluation metrics:

- The % error; the absolute difference in lesion volume (VD) between the output segmentation masks and the manual annotations masks:

$$VD = \frac{|TP_{auto} - TP_{gt}|}{TP_{gt}} \times 100 \quad (1)$$

where $TP_{auto}$and TP_{gt}represent the segmented voxels in the output and manual annotations masks.

- Sensitivity; the True Positive Rate (TPR) between the output segmentation and the manual lesion annotations masks:

$$TPR = \frac{TP}{(TP + FN)} \times 100 \quad (2)$$

in which TP and FN represent sequentially, the number of correct and missed lesion per region.

- False discovery rate; the False Positive Rate (FPR) between output segmentation masks and manual lesion annotations:

$$FPR = \frac{FP}{(FP + TP)} \times 100 \quad (3)$$

where FP represents lesion region candidates numbers incorrectly classified as a lesion.

5.3 Experiment details

The T1-w image was rigidly co-registered to the standard Montreal Neurological Institute (MNI) atlas. Both of FLAIR and T2-w image modalities had been rigidly co-registered to the T1-w space. Therefore, the proposed method began with skull stripping using the brain extraction tools BET in the MRIcro software [38] and then intensity correction was applied using N3 [40]. All training and testing images had been supplied with an isotropic resolution of 0.5 × 0.5 × 0.5 mm.

The training of the proposed method on the available 20 images provided with their manual expert annotations. It is adjusted the maximum number of the training epochs at 100 with early stopping at 15 to every network.

5.3.1 MICCAI2008 Public Dataset (Training Set)

The measures were calculated utilizing the expert label map given in a dataset. In terms of DSC, the proposed method using FLAIR modality achieved the highest average score of 25% and 29% for UNC and CHB datasets, respectively, compared to the other modalities. T2-w was better than T1-w with an average DSC score of 17% and 27% for UNC and CHB, respectively. According to these results shown in Table 1, the proposed R-CNN gave the best results using FLAIR, then T2-w while T1-w modality gave the lowest scores. Therefore, only FLAIR images were estimated for the test.

5.3.2 MICCAI2008 Private Dataset (Testing Set)

The set of known metrics defined in [32] was used for a quantitative evaluation performed on the private data set. The organizers of the challenge had defined the average score to allow a quantitative comparison between approaches and human exports. If an independent human observer would perform manual segmentation, each metric is related to the result that could be expected. Therefore, 100 points mean that it is a perfect result and for an independent human observer, a predefined amount of 90 for a score is typical. The algorithm results on the MICCAI2008 private dataset is an average score of 72%. The comparison between the proposed method and other methods using the MICCAI challenge test data set is shown in Table 2. The performance of the proposed system is superior to the method by Mechrez et al. [48], but not as well as Geremia et al. [33], Roura et al. [9], and Shiee et al. [49]. Figure 2 shows some examples of the proposed MS lesion segmentation results.

6. CONCLUSION

An automated MS segmentation method was presented to delineate lesions in MRI sequences based on the R-CNN model. The lesion of MS can be of any size, shape or position in the entire brain. As a result, traditional R-CNN is not adaptable for accurate segmentation and classification of MS lesions. Therefore, an R-CNN patch technique is proposed for segmenting the MS lesion. In future work, it is proposed to use a deeper network, which can potentially increase performance. Moreover, adding more experiential results into a larger dataset. Finally, adding a false positive reduction step utilizing another CNN to classify the lesion from the background.

Table 1. The results of the proposed patch-wise R-CNN method tested individually on FLAIR, T2 and T1 modalities of the MS challenge training set using LOSO cross-validation.

	Flair			**T2**			**T1**		
Patient	TRP	PPV	DSC	TRP	PPV	DSC	TRP	PPV	DSC
UNC 01	43	3	6	22	2	4	10	1	2
UNC 02	79	21	33	41	14	21	25	12	16
UNC 03	40	26	32	20	17	18	14	14	14
UNC 04	75	29	42	47	17	25	28	13	18
UNC 05	72	10	18	65	6	10	55	7	12
UNC 06	22	6	9	17	5	8	9	4	6
UNC 07	51	7	12	79	7	14	27	2	4
UNC 08	65	20	30	49	15	23	47	16	24
UNC 09	74	24	36	53	19	28	37	14	20
UNC 10	59	25	36	37	14	20	20	11	14
Average	58	17	25	43	12	17	27	9	13
CHB 01	59	28	38	25	19	21	20	24	22
CHB 02	62	18	27	46	19	27	60	15	26
CHB 03	77	13	22	59	16	25	38	19	23
CHB 04	78	9	16	37	16	22	38	11	17
CHB 05	56	16	24	48	22	30	30	14	19
CHB 06	41	31	35	30	23	26	27	20	21
CHB 07	23	29	26	42	30	35	30	21	25

Table 2. Comparison among our proposed R-CNN technique and the more advanced methods in the MS challenge test set

Method	**UNC Rater**			**CHB Rater**			**Score**
	VD	TPR	FPR	VD	TPR	FPR	
Roura et al. [9]	65.2	44.9	43.2	158.9	55.4	40.5	82.34
Geremia et al. [33]	42.1	43.5	76.7	52.4	59.0	71.5	82.07
Shiee et al. [49]	63.7	49.0	74.9	85.2	55.8	70.5	79.90
Proposed method	267.9	63.4	85.8	238.0	70.2	86.2	72
Mechrez *et al.* [48]	119.4	20.1	72.9	73.2	21.1	73.0	71.83

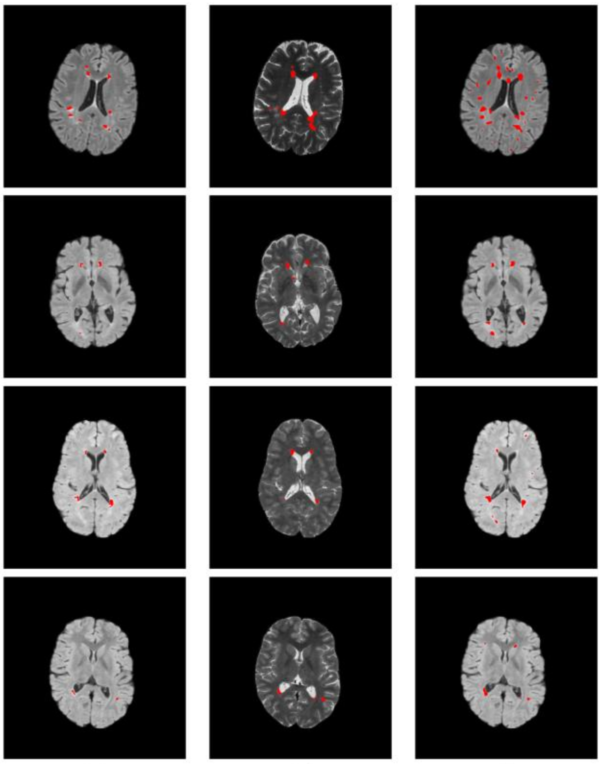


Figure 2: MS lesion segmentation results. (a) Ground-truth. R-CNN results using T2 images (b) and FLAIR images (c).

REFERENCES

[1] A. Compston and A. Coles, "Multiple sclerosis," Lancet, vol. 372, no. 9648, pp. 1502–1517, Oct. 2008

[2] Tomas-Fernandez, X., Warfield, S. K., 2015. A model of population and subject (MOPS) intensities with application to multiple sclerosis lesion segmentation. IEEE Trans. Med. Imag. 34 (6), 1349{1361.

[3] Mohaddeseh Peyvandi and Ali A. Pouyan, 2015, " Automatic Segmentation of Multiple Sclerosis Lesions in Brain MR Images", DOI: 10.14738/jbemi.25.1560

[4] D. Garc´ıa-Lorenzo, S. Francis, S. Narayanan, D. L. Arnold, and D. L. Collins, "Review of automatic segmentation methods of multiple sclerosis white matter lesions on conventional magnetic resonance imaging." Med. Image Anal., vol. 17, no. 1, pp. 1–18, Jan. 2013.

[5] L. P. Clarke, R. P. Velthuizen, M. A. Camacho, J. J. Heine, M. Vaidyanathan, L. O. Hall, R. W. Thatcher, and M. L. Silbiger, "MRI segmentation: methods and applications." Magn. Reson. Imaging, vol. 13, no. 3, pp. 343–68, Jan. 1995

[6] M. Filippi, M. A. Horsfield, S. Bressi, V. Martinelli, C. Baratti, P. Reganati, A. Campi, D. H. Miller, and G. Comi, "Intra- and inter-observer agreement of brain MRI lesion volume measurements in multiple sclerosis. A comparison of techniques." et al. vol. 118 (Pt 6, pp. 1593– 600, Dec. 1995.

[7] Lladó, X., Oliver, A., Cabezas, M., Freixenet, J., Vilanova, J.C., Quiles, A., Valls, L., Ramió-Torrentà, L., Rovira, À., 2012. Segmentation of multiple sclerosis lesions in brain MRI: a review of automated approaches. Inf. Sci. 186(1), 164–185. http://dx.doi.org/10.1016/j.ins.2011.10.011

[8] D. Garc´ıa-Lorenzo, S. Prsegmentation inold, D. L. Collins, C. Barillot, Trimmed-likelihood estimation for focal lesions and tissue segmentation in multisequence mri for multiple sclerosis, Medical Imaging, IEEE Transactions on 30 (8) (2011) 1455–1467.

[9] Eloy Roura, Arnau Oliver, Mariano Cabezas, Sergi Valverde, Deborah Pareto, Joan C Vilanova, Llu´ıs Ramio-Torrent ´ a, ` Alex Rovira, and ` Xavier Llado. A toolbox for multiple sclerosis lesion segmentation. ´ Neuroradiology, 57(10):1031–1043, 2015

[10] Strumia, M., Schmidt, F. R., Anastasopoulos, C., Granziera, C., Krueger, G., Brox, T., 2016. White matter MS-lesion segmentation using a geometric brain model. IEEE Trans. Med. Imag. 35 (2), 1636{1646

[11] Freire, P. G., Ferrari, R. J., 2016. Automatic iterative segmentation of multiple sclerosis lesions using Student's t mixture models and probabilistic anatomical atlases in FLAIR images. Computers in Biology and Medicine 73, 10{23.

[12] García-Lorenzo, D., Prima, S., Collins, D.L., Arnold, D.L., Morrissey, S.P., Barillot, C., 2008a. Combining robust expectation maximization and mean shift algorithms for multiple sclerosis brain segmentation. MICCAI Workshop on Medical Image Analysis on Multiple Sclerosis (Validation and Methodological Issues) (MIAMS'2008) 82–91.

[13] Harmouche, R., Subbanna, N. K., Collins, D. L., Arnold, D. L., Arbel, T., 2015. Probabilistic multiple sclerosis lesion classification based on modeling regional intensity variability and local neighborhood information. IEEE Trans. Biomed. Engg. 62 (5), 1281, 1292.

[14] M. L. Seghier, A. Ramlackhansingh, J. Crinion, A. P. Leff, C. J. Price, Lesion identification using unified segmentation-normalization models and fuzzy clustering, Neuroimage 41 (4) (2008) 1253–1266.

[15] Carass, A., et. al., 2017. Longitudinal multiple sclerosis lesion segmentation: resource & challenge. NeuroImage 148, 77{102.

[16] Sweeney, E. M., Shinohara, R. T., Shiee, N., Mateen, F. J., Chudgar, A. A., Cuzzocreo, J. L., Calabresi, P. A., Pham, D. L., Reich, D. S., 2013. OASIS is automated statistical inference for segmentation, with applications to multiple sclerosis lesion segmentationin MRI. NeuroImage: Clinical 2, 402413

[17] Lao, Z., Shen, D., Liu, D., Jawad, A. F., Melhem, E. R., Launer, L. J., Bryan, R. N.,

Davatzikos, C., 2008. Computer-assisted segmentation of white matter lesions in 3D MR images, using support vector machine. Academic Radiology 15 (3), 300{313.

[18] M. D. Steenwijk, P. J. Pouwels, M. Daams, J. W. van Dalen, M. W. Caan, E. Richard, F. Barkhof, H. Vrenken, Accurate white matter lesion segmentation by k nearest neighbor classification with tissue type priors (knn-ttps), NeuroImage: Clinical 3 (2013) 462–469. A. Cerasa, E. Bilotta, A. Augimeri, A. Cherubini, P. Pantano, G. Zito, P. Lanza, P. Valentino, M. C. Gioia, A. Quattrone, A cellular neural network methodology for the automated segmentation of multiple sclerosis lesions, Journal of neuroscience methods 203 (1) (2012) 193–199.

[19] Y. Xie, X. Tao, White matter lesion segmentation using machine learning and weakly labeled mr images, in: SPIE Medical Imaging, International Society for Optics and Photonics, 2011, pp. 79622G–79622G

[20] Roy, S., He, Q., Carass, A., Jog, A., Cuzzocreo, J. L., Reich, D. S., Prince, J. L., Pham, D. L., 2014. Example based lesion segmentation. In: Proceedings of SPIE Medical Imaging (SPIE). Vol. 9034.p. 90341Y.

[21] Jog, A., Carass, A., Pham, D. L., Prince, J. L., 2015. Multi-output decision trees for lesion segmentation in multiple sclerosis. In: Proceedings of SPIE Medical Imaging (SPIE). Vol. 9413. p. 94131C

[22] Griffanti, L., Zamboni, G., Khan, A., Li, L., Bonifacio, G., Sundaresan, V., Schulz, U. G., Kuker, W., Battaglini, M., Rothwell, P. M., 2016. BIANCA (Brain Intensity Ab Normality Classification Algorithm): A new tool for automated segmentation of white matter hyperintensities. NeuroImage 141, 191205

[23] Roy, S., Carass, A., Prince, J. L., Pham, D. L., 2015a. Longitudinal patch-based segmentation of multiple sclerosis white matter lesions. In: Machine Learning in Medical Imaging. Vol. 9352. pp. 194{202.

[24] Guizard, N., Coupe, P., Fonov, V. S., Manjon, J. V., Arnold, D. L., Collins, D. L., 2015. Rotation-invariant multi-contrast non-local means for MS lesion segmentation. NeuroImage: Clinical 8, 376{ 389.

[25] Deshpande, H., Maurel, P., Barillot, C., 2015. Adaptive dictionary learning for competitive classification of multiple sclerosis lesions. In: Intl. Symp. on Biomed. Imag. (ISBI). pp. 136{139.

[26] D Ciresan, Alessandro Giusti, and J Schmidhuber. Deep neural networks segment neuronal membranes in electron microscopy images. Advances in Neural Information Processing Systems, pages 1–9, 2012.

[27] LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436444.

[28] Moeskops, P., Viergever, M. A., Mendrik, A. M., de Vries, L. S., Benders, M. J. N. L., Isgum, I., 2016a. Automatic segmentation of MR brain images with a convolutional neural network. IEEE Trans. Med. Imag. 35 (5), 1252{1261.

[29] Kamnitsas, K., Ledig, C., Newcombe, V. F., Simpson, J. P., Kane, A. D., Menon, D. K., Rueckert, D., Glocker, B., 2017. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Med. Image Anal. 36, 61{78.

[30] Havaei, M., Guizard, N., Chapados, N., Bengio, Y., 2016. HeMIS: Hetero-modal image segmentation. In: Med. Image Comp. and Comp. Asst. Intervention (MICCAI). pp. 469- 477.

[31] Styner, Martin, Lee, Joohwi, Chin, Brian, et al., 2008. 3D segmentation in the clinic: A grand challenge II: MS lesion segmentation. Midas Journal 2008, 1–6

[32] Geremia, Ezequiel, Menze, Bjoern H., Clatz, Olivier, et al., 2010. Spatial decision forests for MS lesion segmentation in multi-channel MR images. In: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). Springer, pp. 111–118.

[33] Brosch, Tom, Tang, Lisa Y.W., Yoo, Youngjin, et al., 2016. Deep 3D convolutional encoder networks with shortcuts for multiscale feature integration applied to multiple sclerosis lesion segmentation. IEEE Transactions on Medical Imaging 35 (5), 1229–1239.

[34] Roth, H. R., Lu, L., Liu, J., Yao, J., Seff, A., Cherry, K., Kim, L., Summers, R. M., 2016. Improving computer-aided detection using convolutional neural networks and random view aggregation. IEEE Trans. Med. Imag. 35 (5), 1170{1181.

[35] M. Maleki, M. Teshnehlab, M. Nabavi, Diagnosis of Multiple Sclerosis (MS) Using Convolutional Neural Network (CNN) from MRIs, Global Journal of Medicinal Plant Research 1 (1) (2012) 50–54.

[36] Prieto, J. C., Cavallari, M., Palotai, M., Pinzon, A. M., Egorova, S., Styner, M., Guttmann, C. R. G., 2017. Large deep neural networks for MS lesion segmentation. In: Proceedings of SPIE Medical Imaging (SPIE). Vol. 10133. p. 10133F.

[37] Smith, SM., Fast robust automated brain extraction. Human Brain Map, 2002. 17(3): p. 143–155.

[38] A. Simmons, P. S. Tofts, G. J. Barker, S. R. Arridge, Sources of intensity nonuniformity in spin echo images at 1.5 T, Magnetic Resonance in Medicine 32 (1) (1994) 121–128.

[39] Led, JG., AP. Zijdenbos and, AC. Evans, A nonparametric method for automatic correction of intensity nonuniformity in MRI data. Medical Imaging, IEEE Transactions on, 1998. 18(1): p. 87–97.

[40] Perona, P. and, J. Malik, Scale-space edge detection using anisotropic diffusion. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 1990. 12(7): p. 629-639.

[41] Y. LeCun et al., “ Backpropagation applied to handwritten zip code recognition,” Neural Comput., vol. 1, no. 4, pp. 541–551, 1989.

[42] He K, Zhang X, Ren S, Sun J: “Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification,” in 2015 I.E. International Conference on Computer Vision (ICCV), 2015.

[43] Girshick, R., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: Computer Vision and Pattern Recognition (CVPR), pp. 580–587 (2014).

[44] Girshickv, R.: Fast R-CNN. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 1440–1448 (2015).

[45] Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: towards real-time object detection with region proposal networks. In: NIPS (2015)

[46] Koen EA Van de Sande, Jasper RR Uijlings, Theo Gevers, and Arnold WM Smeulders, “Segmentation as selective search for object recognition,” in Computer Vision (ICCV), 2011 IEEEInternational Conference on. IEEE, 2011, pp. 1879–1886.

[47] Mechrez R, Goldberger J, Greenspan H. 2016. Patch-Based Segmentation with Spatial Consistency: Application to MS Lesions in Brain MRI. International Journal of Biomedical Imaging. 2016.

[48] Shiee N, Bazin PL, Ozturk A, Reich DS, Calabresi PA, Pham DL. 2010. A topology-preserving approach to the segmentation of brain images with multiple sclerosis lesions. NeuroImage. 49(2):1524-1535.