Journal of Medical Imaging

MedicalImaging.SPIEDigitalLibrary.org

# BrainSegNet: a convolutional neural network architecture for automated segmentation of human brain structures

Raghav Mehta
Aabhas Majumdar
Jayanthi Sivaswamy

SPIE.

Raghav Mehta, Aabhas Majumdar, Jayanthi Sivaswamy, "BrainSegNet: a convolutional neural network architecture for automated segmentation of human brain structures," *J. Med. Imag.* **4**(2), 024003 (2017), doi: 10.1117/1.JMI.4.2.024003.

# BrainSegNet: a convolutional neural network architecture for automated segmentation of human brain structures

**Raghav Mehta,* Aabhas Majumdar, and Jayanthi Sivaswamy**
Centre for Visual Information Technology (CVIT), International Institute of Information Technology - Hyderabad (IIIT-H), Hyderabad, India

**Abstract.** Automated segmentation of cortical and noncortical human brain structures has been hitherto approached using nonrigid registration followed by label fusion. We propose an alternative approach for this using a convolutional neural network (CNN) which classifies a voxel into one of many structures. Four different kinds of two-dimensional and three-dimensional intensity patches are extracted for each voxel, providing local and global (context) information to the CNN. The proposed approach is evaluated on five different publicly available datasets which differ in the number of labels per volume. The obtained mean Dice coefficient varied according to the number of labels, for example, it is $0.844 \pm 0.031$ and $0.743 \pm 0.019$ for datasets with the least (32) and the most (134) number of labels, respectively. These figures are marginally better or on par with those obtained with the current state-of-the-art methods on nearly all datasets, at a reduced computational time. The consistently good performance of the proposed method across datasets and no requirement for registration make it attractive for many applications where reduced computational time is necessary.  [DOI: 10.1117/1.JMI.4.2.024003]

Keywords: multiatlas segmentation; brain MRI; convolutional neural networks.

Paper 16218RRR received Sep. 24, 2016; accepted for publication Mar. 28, 2017; published online Apr. 20, 2017.

## 1 Introduction

Quantitative analysis of the neuroimaging data requires cortical and noncortical structural segmentation. Such analysis is critical in many tasks, such as the assessment of several neurodegenerative disorders, development of neonatal brain, etc. Manual labeling of these structures is unsuitable for studies involving large datasets since it is a slow process and prone to human errors. Automatic segmentation addresses these problems. Though there are different methods for automatic segmentation, a popular one is based on the use of multiple atlases, known as multiatlas segmentation (MAS).

The most widely used approach for MAS is based on nonrigid registration and label fusion. A typical nonrigid registration-based method for MAS follows these steps: (i) selection of the relevant source atlases (with labeled voxels) from a training set[1] and their nonrigid registration to the target volume, (ii) label propagation from the source to the target space,[2] and (iii) finally, label fusion[3] to combine the propagated labels into a segmentation estimate for the target volume. The nonrigid registration and the label fusion technique determine the accuracy of these methods. This approach for MAS is known as multiatlas label propagation (MALP). A good survey of different methods for MAS task can be found in Ref. 4.

A key drawback of the MALP methods is the computational cost. The nonrigid registration step can require as long as 2 to 20 h,[5] while the label fusion[6] and atlas selection[7] typically require 2 to 3 h. Furthermore, since the total computational time is linearly proportional to the number of training atlases, it linearly increases with an increase in the number of training atlases.

Alternatively, patch-based techniques[8,9] have also been investigated as a solution for MAS. Here, the labeling of each voxel is done by comparing its surrounding patch with patches in training data in which the labels of the central voxels are known. These patch-based techniques also typically require 2 to 3 h for labeling of new volume.[10]

A reduction in the computational cost is possible with offline learning,[11–16] where the structure segmentation is proposed as a voxel classification problem. Here, a model is learned on the training data using different machine learning algorithms and is used to classify the voxels in an unseen volume in a short time. An example of this is the use of a set of atlas forests (AFs).[11] Here, each AF encodes a single atlas and a probabilistic atlas is constructed by iterative nonrigid groupwise registration of training samples to their mean. During testing, a new volume is first affine registered to the probabilistic atlas followed by a coarse, nonrigid registration, thus leading to computational efficiency.

Spurred by the success of deep learning in computer vision,[17,18] convolutional neural network (CNN)-based techniques have been explored for the segmentation of various anatomical structures, from the pancreas in CT images[19] to neuronal structures in electron microscopic stacks.[20]

In neuroimages, segmentation of images into three basic tissue types, such as gray matter (GM), white matter (WM), and cerebrospinal fluide, in different age groups has been attempted using CNN,[21,22] while segmenting eight different tissue types has been reported in Ref. 23.

*Address all correspondence to: Raghav Mehta, E-mail: raghav.mehta@research.iiit.ac.in

2329-4302/2017/$25.00 © 2017 SPIE

The task of segmenting cortical and subcortical structures is similar to the tissue-labeling task, albeit more challenging as the number of structures of interest is generally large, typically 32 to 134, with the intensity information being inadequate as many structures belong to both GM and WM tissues. Hence, the contextual information for this task is essential to distinguish between structures and correctly identify their location in left and right hemispheres of the brain as shown in Ref. 13. Nonrigid registration (of a labeled atlas) has been a natural choice for solving this task as it has embedded contextual information. This contextual information is obtained with the help of propagated atlas priors[2,24] or alternately the spatial information is provided explicitly as prior probability of structure in patch-based approaches.[8–10,25]

Some recent works take the CNN-based approach for structure segmentation.[13–16] Every voxel is represented in Ref. 13 by a set of features which includes local appearance information and position relative to the centroids of different substructures/segments of the brain as contextual information. Segmenting a new volume involves an iterative two-step solution: one to generate a rough labeling (for centroid information) and the other to refine it. These two steps are repeated until convergence.

A fully convolutional network (FCN)-based CNN approach has also been proposed[14] to segment deep brain structures (typically 10 to 12 structures). Since contextual or three-dimensional (3-D) information is not a part of the input in this method, it is necessary to use a separate Markov random field (MRF) as postprocessing for label consistency. Alternately, Hough voting has also been used as the postprocessing step.[15] Training a separate CNN for each structure of interest has also been proposed for deep brain structure segmentation in Ref. 16. Dynamic random walker with decayed region of interest is then used to enforce label consistency. The use of separate CNNs for each structure makes it impractical for scenarios where it is necessary to segment up to 134 structures. Thus, existing CNN-based solutions for structure segmentation have the following drawbacks: they are iterative in nature, require postprocessing, do not permit end-to-end training, and often require a registration step. Furthermore, most of the existing CNN based methods have addressed largely labeling of the subcortical structures and not the whole brain segmentation.

We propose a noniterative as well as end-to-end trainable CNN-based solution for this task. The salient features of our solution are:

- Innovative patches: The CNN is trained on two-dimensional (2-D) and 3-D patches to capture four types of voxel appearance information which encode local intensity profile as well as global context.
- No requirement for registration: Unlike the existing methods, no registration is required to segment a new volume.
- Comprehensive labeling: A structure segmentation task has been demonstrated for segmentation of both cortical and subcortical structures.
- Good performance: A reduction in computational time by a factor of 70 is achieved relative to nonrigid registration-based approaches; this is independent of the training set size. The proposed solution also gives comparable or marginally better performance than the current state-of-the-art techniques on four out of five public datasets in terms of accuracy.

## 2 Methodology

CNN is a deep learning architecture inspired by biological networks akin to the multilayer perceptron. Basic blocks of a CNN are: (i) a convolutional layer (2-D/3-D) to detect local features at different positions in an image through a set of learnable filters (or kernels), (ii) a maxpooling layer (2-D/3-D), which downsamples the output of a layer thus progressively reducing the number of parameters and computation, (iii) a fully connected (FC) layer, which is an extension of the original multilayer perceptron, (iv) a dropout layer[26] to effectively regularize the network and reduce overfitting to the training data, and (v) activation functions, such as rectified linear unit, tanh, and softmax.[17] The architecture we propose is shown in Fig. 1.

In the structure segmentation task, three types of information are useful for segmentation. (i) Local intensity profile to distinguish between structures belonging to different tissues, (ii) context, necessary to encode the spatial configuration of structures, and (iii) 3-D information for attaining label consistency across slices.

The proposed CNN architecture, henceforth referred to as "BrainSegNet," has this information provided via 2-D/3-D patches of various sizes. Around every voxel, the following patches are extracted: (i) three 2-D orthogonal patches of size $31 \times 31$ voxel, extracted from the sagittal, coronal, and axial ("sca") planes to provide local 2-D intensity profile ($31 \times 31 \times 3$), (ii) a 3-D patch of size $21 \times 21 \times 21$, to provide local 3-D intensity profile ($21 \times 21 \times 21 \times 1$), (iii) three 2-D orthogonal patches of size $93 \times 93$, downsampled by a factor of 3, to provide global information regarding input voxel, from the sca planes, ($31 \times 31 \times 3$), (iv) a 3-D patch of size $63 \times 63 \times 63$, downsampled by a factor of 3, to provide global 3-D information ($21 \times 21 \times 21 \times 1$). Sample patches for seven different voxels are shown in Fig. 2.

Inspired by the VGGnet,[27] small-sized kernels were chosen for the convolutional layers in our architecture, as shown in Fig. 1. Each of the four type of patches mentioned above is considered as an input in our network and has a separate processing pipeline for them. Each branch has a cascade of two convolutional layers followed by one maxpooling layer. Each convolutional layer has $3 \times 3(2-\mathrm{D})/3 \times 3 \times 3(3-\mathrm{D})$ kernel size, while the numbers of filters are 8, 16, 24, and 32, respectively. The maxpooling uses $2 \times 2(2-\mathrm{D})/2 \times 2 \times 2(3-\mathrm{D})$ filters with a stride equal to 2, thus reducing the patch size by half.

The output of all the four branches is flattened and concatenated to form a single one-dimensional array. This is passed through two FC layers with 2000 and 300 neurons, respectively, and softmax layers ($N$ neurons, $N$ = total number of desired labels including background) in a sequential manner. A dropout layer (with probability 0.5) is applied between the two FC layers as well as between the last FC layer and the softmax layer (in blue).

A weighted categorical cross entropy function (with respect to classes) is used to handle the class-imbalance problem. This loss function and the weights are defined such that the weight increases whenever there are fewer voxels in a particular class

$$L^i = -\sum_n \sum_l w_l * t^i_{n,l} \log\, p^i_{n,l} \quad \text{where,}\ w_l = \frac{\sum_{k=0}^{k=N} m_k}{m_l},$$

where $L^i$ is the loss for volume $i$, $t^i_{n,l}$ is 1 if the true label of voxel $n$ of volume $i$ is $l$ otherwise it is 0, $p^i_{n,l}$ is the probability that the CNN will predict label $l$ for voxel $n$ of volume $i$, $w_l$ denotes the

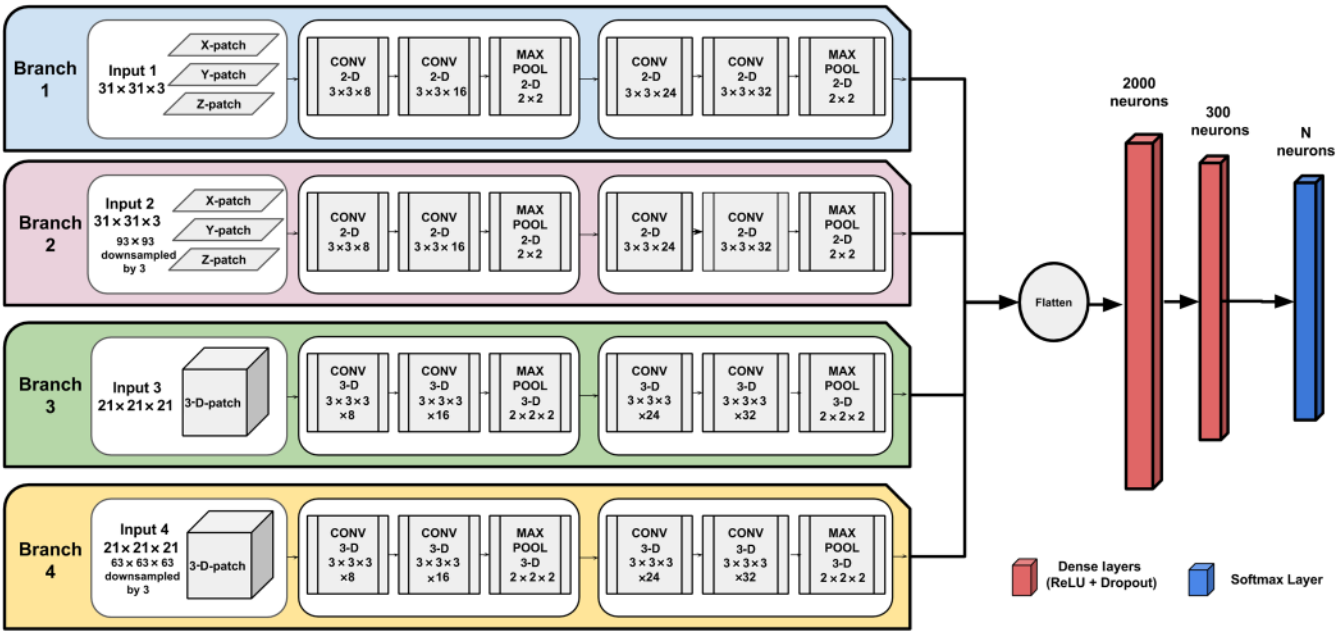


**Fig. 1** Schematic overview of the proposed CNN architecture. The number of neurons $N$ is same as the number of manually marked structures in a dataset (including background).

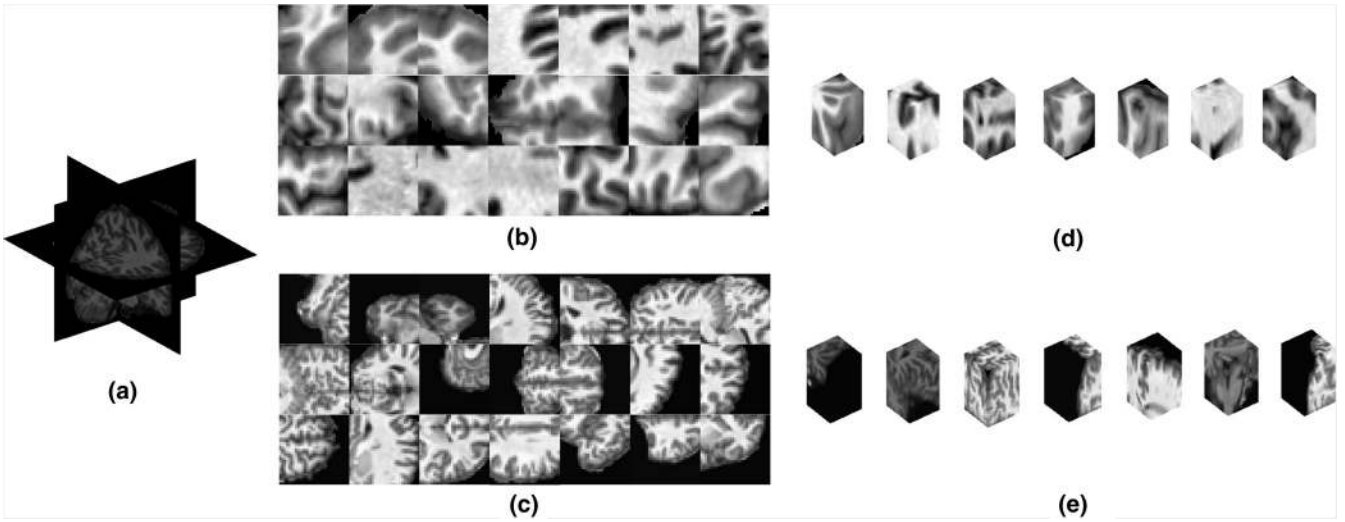


**Fig. 2** Sample input patches: (a) 2.5-D representation of the brain MRI volume. For seven different voxels, (b) the branch 1 ($31 \times 31 \times 3$), (c) branch 2 ($93/3 \times 93/3 \times 3$), (d) branch 3 ($21 \times 21 \times 21$), and (e) branch 4 ($63/3 \times 63/3 \times 63/3$) patches/cubes are also shown. The ordering for (b) and (c) are: coronal (top row), sagittal (middle row), and axial (bottom row) slices. For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.

weight for the $l$'th class, and $m_l$ is the number of voxels of $l$'th class in the training dataset.

### 2.1 *Different Variants of the Proposed Convolutional Neural Network*

As mentioned above, the BrainSegNet has four types of inputs (2-D/3-D local/global) which contribute to the CNN architecture. Analysis of their relative contribution is possible by considering various architectures.

In general, it can be observed that the variation in interstructure intensity of noncortical regions is greater than their spatial positions. This is contrary to the cortical structures, where spatial position (contextual information) varies more than their local intensity information. Thus, the segmentation accuracy should vary with the type of input that is available in the CNN. These observations were verified by considering CNN variants with only one type of input. CNN1: a simple architecture which has only 2-D local information as input. Specifically, an input of size $31 \times 31 \times 3$ is derived from a $31 \times 31$ voxels patch from the three sca planes. CNN2a: a network with only 2-D "global" information as input. The input is now a patch of size $21 \times 21 \times 3$ derived from $63 \times 63$ patches from three sca planes after downsampling by 3. CNN2b: here, the

size of global context considered is wider. In place of the $63 \times 63$ patch, a $93 \times 93$ patch is used to derive the 2-D global information.

One can say that a combination of both local and global information is beneficial for the accurate segmentation of both cortical and noncortical structures. To observe the performance of this combination of inputs, we considered a variant of the CNN architecture (CNN3). Here, the inputs are both 2-D and 3-D. Specifically, the input consists of 2-D local information from sca planes ($31 \times 31 \times 3$), 2-D global information from sca planes ($63 \times 63$ downsampled by factor 3, $21 \times 21 \times 3$), and 3-D local ($15 \times 15 \times 15$) information. The 3-D input patches can provide segmentation consistency between successive slices and remove the requirement of postprocessing. The proposed BrainSegNet has all the necessary 2-D/3-D local/global information as input.

An alternative to using both 2-D and 3-D patches is to use only 3-D inputs as in Ref. 28. However, processing 3-D patches requires greater graphics processing unit (GPU) memory and increases the computational time (for the same size of input) as 3-D convolution is computationally more intensive than 2-D. Hence, this variant was not considered. All four variants and the BrainSegNet were tested on a public dataset and the results are presented in Sec. 4.1.

## 3 Dataset

BrainSegNet was evaluated on five different publicly available datasets, with varying numbers of structures per volume, as described in Table 1. As indicated in the table, except for the MICCAI-2012, every dataset is split randomly into two equal-sized training and testing sets.

### 3.1 *MICCAI-2012 Dataset*

This dataset was released as a part of a workshop on multiatlas labeling in MICCAI-2012 (Ref. 29). It consists of 15 training images and 20 testing images from the OASIS[30] project. A detailed description of the acquisition parameters can be found on the OASIS website. All images have 134 manually segmented structures provided by Neuromorphometrics, Inc. The large number of marked structures makes it a challenging dataset for structure segmentation task.

### 3.2 *International Brain Segmentation Repository Dataset*

The International Brain Segmentation Repository (IBSR) dataset has 18, 3-D T1-weighted MR images of 1.5-mm-thick cortical slices. Manual segmentation of 32 structures (primarily noncortical) is provided by the Center for Morphometric Analysis at Massachusetts General Hospital.

**Table 1** Dataset description.

| Dataset | No. of structures | No. of atlases | Train volumes | Test volumes |
|---|---|---|---|---|
| MICCAI-2012 | 134 | 35 | 15 | 20 |
| IBSR | 32 | 18 | 9 | 9 |
| LPBA40 | 54 | 40 | 20 | 20 |
| Hammers67n20 | 67 | 20 | 10 | 10 |
| Hammers83n30 | 83 | 30 | 15 | 15 |

### 3.3 *LONI-LPBA40 Dataset*

The LONI-LPBA40 dataset consists of T1-weighted MRI scans of 40 healthy volunteers. A total of 50 cortical and 4 subcortical structures along with the brainstem and the cerebellum are delineated by trained raters for each volume. Details of acquisition parameters can be found in Ref. 31. In our work, the brainstem and cerebellum were excluded from assessment, as they were removed by the skull-stripping step (Ref. 32).

### 3.4 *IXI Datasets*

Hammers67n20 and Hammers83n30 are two sets, consisting of 20 and 30 T1-weighted MR image data, provided as a part of the IXI database (Ref. 33). Details of the acquisition parameters can be found in Ref. 2. Hammers83n30 has much more detailed segmentation of the gyrus in the frontal and temporal lobs compared to Hammers67n20. This is clearly visible in Figs. 3(d) and 3(e).

## 4 Experiments and Results

All the volumes were preprocessed as follows: intensity inhomogeneity correction was done using N4-bias correction algorithm[34] followed by skull stripping[35] [Brain Extraction Tool (BET)] (Ref. 36) and intensity normalization by subtracting mean intensity of a volume and dividing by the standard deviation of a volume.

BrainSegNet was trained on NVIDIA K40 GPU, with 12 GB of RAM for 30 epochs using stochastic gradient descent with momentum 0.75 and learning rate 0.05. Learning rate was reduced by half at every 10 epochs. Two million patches were extracted from each training volume. The training time was roughly 2 days. The code was written in Python using Keras library. A new test volume can be segmented in 15 to 20 min, which is a 60 to 70 fold reduction compared to standard nonrigid registration-based methods, which require 20 to 25 h for segmentation.

The segmentation performance was quantitatively assessed using the mean Dice coefficient (DC), which is defined as follows. Let $A$ and $B$ denote the binary segmentation labels generated manually and computationally, respectively. The DC is defined as

$$\mathrm{DC}(A,B) = \frac{2|AB|}{|A| + |B|},$$

where $|A|$ denotes the number of positive elements in the binary segmentation A and $|AB|$ is the number of shared positive elements by A and B. $\mathrm{DC} \in [0,1]$. A higher DC value indicates a better segmentation performance.

Sample axial slices drawn from five datasets, their manual segmentation and output of the BrainSegNet are shown in Fig. 3. It should be noted that a "smooth" segmentation of cortical and subcortical structures is obtained without any postprocessing. This is due to the fact that the input patches are rich in information, both local intensity as well as global context is provided.

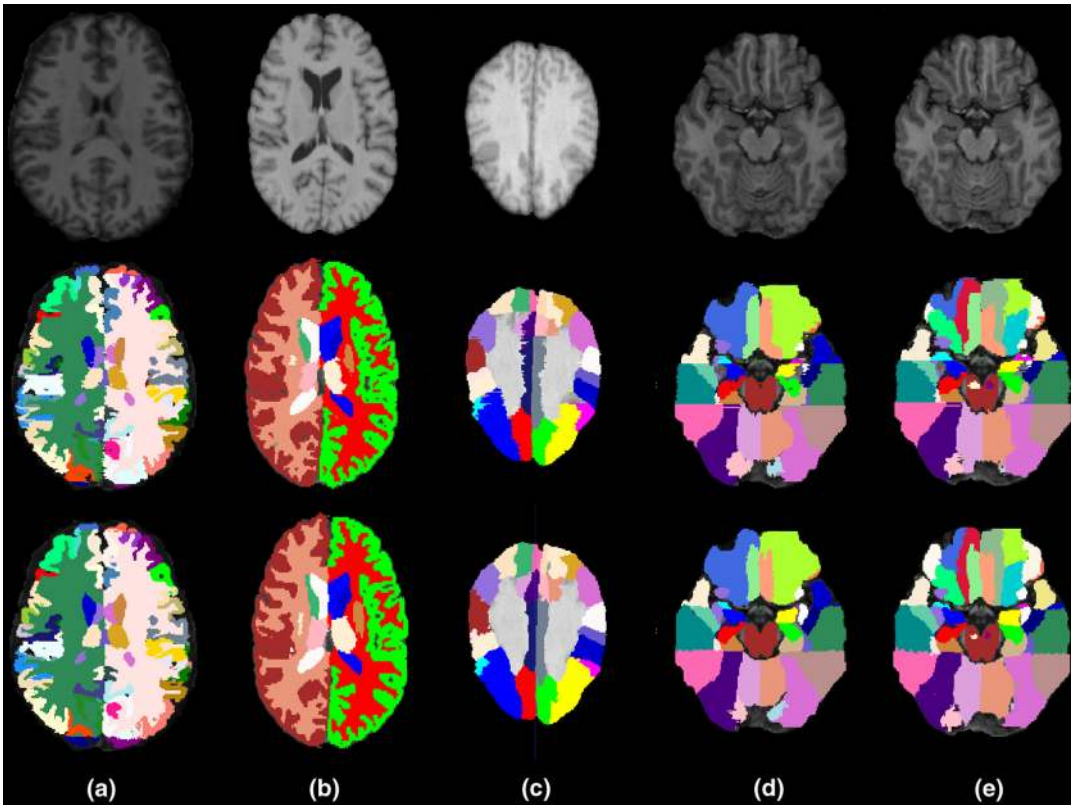


**Fig. 3** Sample images from the different datasets (top row), their manual segmentation (middle row), and output of the BrainSegNet (bottom row). (a) MICCAI-2012, (b) IBSR, (c) LONI-LPBA40, (d) Hammers67n20, and (e) Hammers83n30.

**Table 2** Performance (mean DC) comparison of the BrainSegNet with the various methods (MALP based, patch based, and classification based) for different datasets.

| | Various segmentation methods | | | | | BrainSegNet |
|---|---|---|---|---|---|---|
| MICCAI-2012 | ***0.764***[5] | 0.758[5] | 0.711[5] | 0.737[5] | 0.7275[11] | 0.743* |
| IBSR | *0.835*[10] | | *0.835*[11] | | | **0.844** |
| LONI-LPBA40 | 0.783[8] | 0.784[12] | 0.799[9] | *0.814*[24] | 0.801[11] | **0.824**$^{*,\dagger}$ |
| Hammers67n20 | *0.836*[2] | | 0.754[2] | | | **0.840** |
| Hammers83n30 | *0.801*[33] | 0.752[33] | 0.785[33] | 0.754[33] | 0.789[33] | **0.808**$^{\dagger}$ |

*Statistically significant difference ($p < 0.01$) between the proposed and random forest-based method.[11]
†Statistical significant difference between the proposed and the state-of-the-art method (italics) for respective datasets.
Note: Bold values signify the best performing methods.

A comparison of the performance of the BrainSegNet with the state-of-the-art standard methods for MALP is given in Table 2. This table shows that the performance of the BrainSegNet is comparable or marginally better than other methods, on all five datasets except one.

Since the methods taken up for comparison are computationally expensive, the tabulated results for other methods are drawn from the respective papers without reimplementing them. Hence, to establish the statistical significance ($p$ value) of the results, the one sample $t$-test/$z$-test has been used.

A detailed evaluation of all the variants listed in Sec. 2.1 was done on MICCAI-2012 dataset and results are provided in Sec. 4.1. A detailed evaluation of the proposed method on all the other datasets is provided in Sec. 4.2.

### 4.1 *Performance of Different Variants of the Proposed Convolutional Neural Network on MICCAI-2012*

The CNN variants were evaluated on the MICCAI-2012 dataset as it has the greatest number of labeled structures (134). It should be noted that all these architectures were trained using the same number of patches and optimization parameters as mentioned in Sec. 4.

**Table 3** Mean DC values for different variants of the proposed CNN architecture on MICCAI-2012 dataset.

| | Cortical structures | Noncortical structures | Overall |
|---|---|---|---|
| CNN1 | $0.6306 \pm 0.023$ | $0.7701 \pm 0.024$ | $0.6685 \pm 0.021$ |
| CNN2a | $0.6370 \pm 0.011$ | $0.7535 \pm 0.024$ | $0.6683 \pm 0.010$ |
| CNN2b | $0.6576 \pm 0.011$ | $0.7604 \pm 0.028$ | $0.6852 \pm 0.012$ |
| CNN3 | $0.6758 \pm 0.013$ | $0.7793 \pm 0.022$ | $0.7036 \pm 0.011$ |
| BrainSegNet | $\mathbf{0.7204 \pm 0.012}$ | $\mathbf{0.8053 \pm 0.028}$ | $\mathbf{0.7432 \pm 0.019}$ |

Note: Bold values signify the best performing methods.

The mean DC for the cortical and noncortical structures for different variants is listed in Table 3. The box plots for different noncortical and cortical structures, for these variants, are shown in Figs. 4 and 5, respectively.

CNN1 is expected to give a superior performance for noncortical structures rather than cortical structures and vice versa for CNN2a, as CNN1 encodes local information necessary for noncortical structures while CNN2a encodes global context necessary for cortical structures. This is borne out to be true as CNN1 has better DC (cortical: $0.6306 \pm 0.023$ and noncortical: $0.7701 \pm 0.024$) than CNN2a (cortical: $0.6370 \pm 0.011$ and noncortical: $0.7535 \pm 0.024$) for noncortical structures and vice versa for cortical structures.

Similarly, it can also be expected that an increase in the global context should increase the DC. This is also true as CNN2b ($0.6852 \pm 0.012$) has better overall DC than CNN2a ($0.6683 \pm 0.010$), underscoring the beneficial effect of context in cortical structure labeling. However, the improvement in DC for cortical (CNN2a: $0.6370 \pm 0.011$ and CNN2b: $0.6576 \pm 0.011$) versus noncortical (CNN2a: $0.7535 \pm 0.024$ and CNN2b: $0.7604 \pm 0.028$) structures differs with only a marginal increase for the latter class, which is to be expected.

Since CNN3 has a combination of 2-D local/global and local 3-D information as input, its performance should be better than the variants with only one type of input. The results in Table 3 affirm this, as an improved labeling performance is evident for both cortical ($0.6758 \pm 0.013$) and noncortical ($0.7793 \pm 0.022$) structures relative to CNN1 and CNN2.

Likewise, the performance of the proposed BrainSegNet should be superior to that of all these variants as it has all four global/local 2-D/3-D information as input. This also holds, as the BrainSegNet has the best performance among all variants, with a mean DC of $0.7432 \pm 0.019$ for all structures in the MICCAI-2012 dataset. This is also better than the AF-based method,[11] which reports a mean DC of $0.7275 \pm 0.070$ ($p < 0.01$).

The ranked results of 25 methods, evaluated during the MICCAI-2012 challenge, are reported in Ref. 5. The obtained performance of BrainSegNet places it at rank 5, with the difference in DC value between the best (0.764) and BrainSegNet ($0.7432 \pm 0.019$) being 0.02. This is noteworthy as all the methods in Ref. 5 are based on (linear/nonlinear) registration and none use a CNN-based solution. It should be noted that this ranking is from 2012 and continues to be valid as of now, since most other methods for brain structure segmentation (including CNN based) in literature have not reported results on this dataset. This perhaps is due to the large number (134) of labeled structures in this dataset.

### 4.2 *Evaluation and Comparison of BrainSegNet on Multiple Datasets*

In addition to MICCAI-2012, BrainSegNet is also evaluated on four other public datasets. We present the performance figures

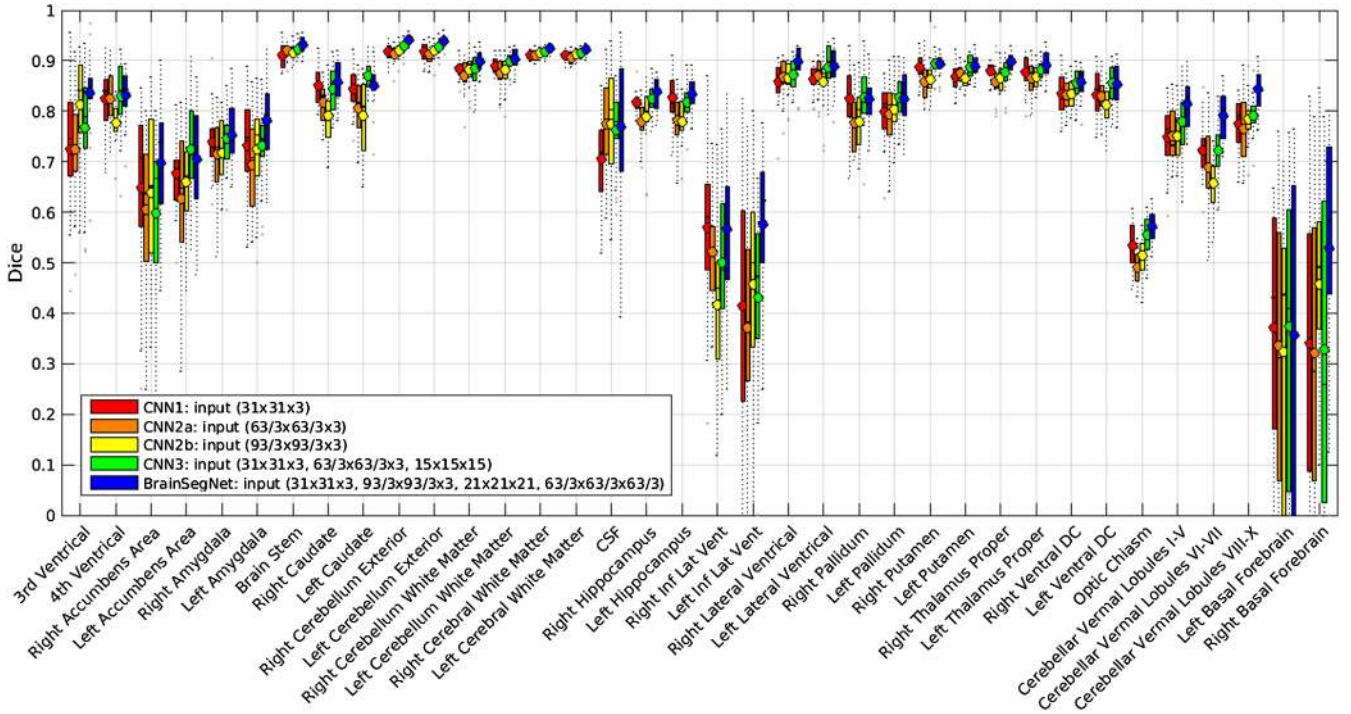


**Fig. 4** Results of labeling of noncortical structures on test volumes of MICCAI-2012 dataset. For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.

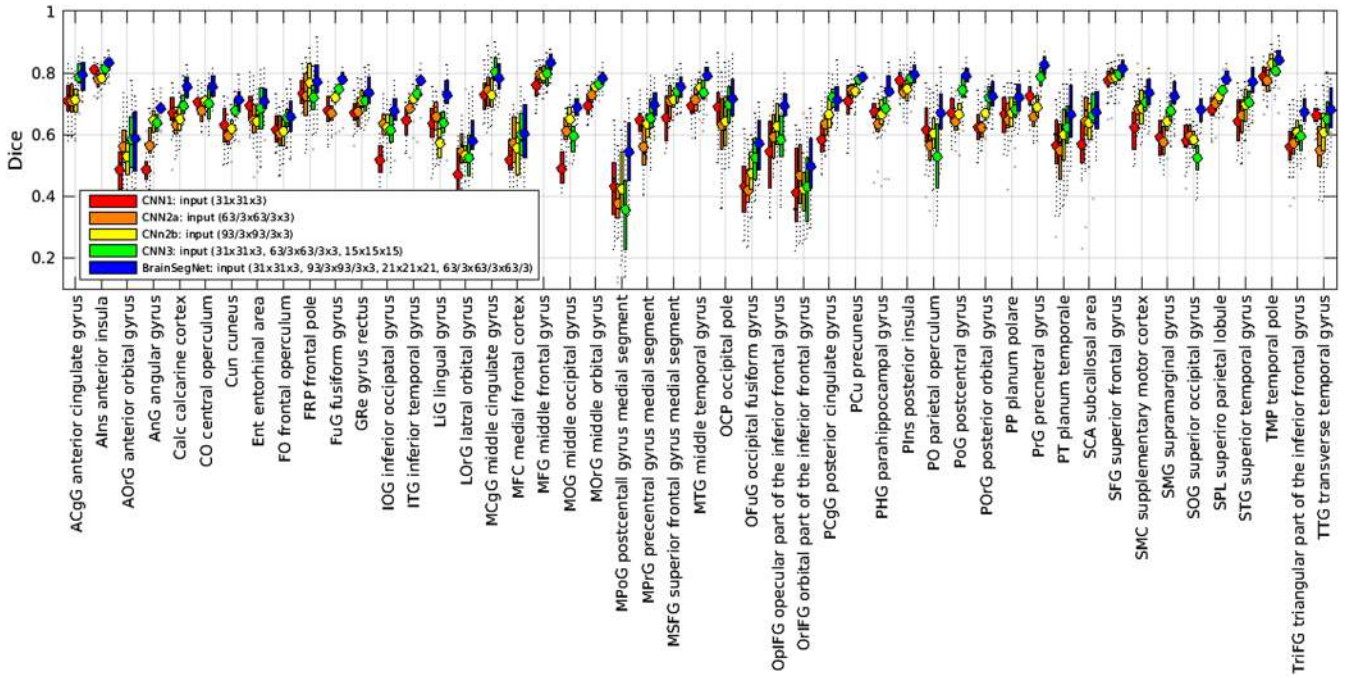


**Fig. 5** Results of labeling of cortical structures on test volumes of MICCAI-2012 dataset (left and right labels shown jointly). For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.

**Table 4** Quantitative comparison on IBSR dataset. Mean DC values are listed for noncortical structures.

| | Ref. 12 | Ref. 10 | BrainSegNet |
|---|---|---|---|
| L. lateral ventrical | 0.85 | **0.93** | $\mathbf{0.93 \pm 0.039}^{*}$ |
| L. thalamus | 0.88 | **0.89** | $0.88 \pm 0.050$ |
| L. caudate | 0.83 | **0.88** | $0.86 \pm 0.047$ |
| L. putamen | 0.84 | 0.89 | $\mathbf{0.91 \pm 0.022}^{*,\dagger}$ |
| L. pallidum | 0.74 | 0.79 | $\mathbf{0.81 \pm 0.089}^{*}$ |
| L. hippocampus | 0.74 | **0.83** | $0.81 \pm 0.065^{*}$ |
| L. amygdala | 0.68 | 0.75 | $\mathbf{0.76 \pm 0.087}^{*}$ |
| L. ventral DC | 0.81 | **0.82** | $\mathbf{0.82 \pm 0.033}$ |
| Third ventrical | 0.74 | 0.80 | $\mathbf{0.81 \pm 0.079}^{*}$ |
| Fourth ventrical | 0.76 | 0.84 | $\mathbf{0.85 \pm 0.094}^{*}$ |
| R. lateral ventrical | 0.85 | **0.92** | $\mathbf{0.92 \pm 0.026}^{*}$ |
| R. thalamus | 0.87 | 0.89 | $\mathbf{0.90 \pm 0.029}^{*}$ |
| R. caudate | 0.81 | **0.89** | $0.88 \pm 0.048^{*}$ |
| R. putamen | 0.84 | 0.89 | $\mathbf{0.91 \pm 0.023}^{*,\dagger}$ |
| R. pallidum | 0.75 | 0.79 | $\mathbf{0.83 \pm 0.086}^{*,\dagger}$ |
| R. hippocampus | 0.76 | **0.83** | $\mathbf{0.83 \pm 0.071}^{*}$ |
| R. amygdala | 0.66 | **0.75** | $0.71 \pm 0.087$ |
| R. ventral DC | 0.81 | **0.82** | $\mathbf{0.82 \pm 0.051}$ |

*The statistically significance difference ($p < 0.05$) between BrainSegNet and Ref. 12.
†The statistically significance difference ($p < 0.05$) between BrainSegNet and Ref. 10.
Note: Bold values signify the best performing methods.

for the whole brain as well as for the individual structures on each of these datasets. These individual structures are chosen according to the reports in the literature on each datasets. It should be noted that the mean DC values, reported from respective papers, are based on leave-one-out assessment whereas ours is based on half-split assessment.

#### 4.2.1 *IBSR dataset*

This dataset has the fewest labeled structures. The mean DC obtained for this dataset with BrainSegNet is $0.844 \pm 0.031$, which is marginally higher than those reported for the current state-of-the-art methods such as in Ref. 10 (0.835) and Ref. 11 ($0.835 \pm 0.042$, $p = 0.1467$). Table 4 lists the mean DC

**Table 5** Quantitative comparison on IBSR dataset for CNN-based methods. DC values are shown for subcortical structures used by Refs. 14 and 16 for evaluation.

| | Ref. 14 | Ref. 16 | BrainSegNet |
|---|---|---|---|
| Thalamus | 0.87 | 0.89 | $\mathbf{0.89 \pm 0.041}$ |
| Putamen | 0.83 | 0.88 | $\mathbf{0.91 \pm 0.022}^{*,\dagger}$ |
| Caudate | 0.78 | **0.87** | $\mathbf{0.87 \pm 0.047}^{*}$ |
| Pallidum | 0.75 | 0.79 | $\mathbf{0.82 \pm 0.083}^{*,\dagger}$ |
| Hippocampus | 0.77 | 0.81 | $\mathbf{0.82 \pm 0.066}^{*}$ |
| Amygdala | 0.65 | 0.67 | $\mathbf{0.74 \pm 0.089}^{*,\dagger}$ |

*The statistically significance difference ($p < 0.05$) between BrainSegNet and Ref. 14.
†The statistically significance difference ($p < 0.05$) between BrainSegNet and Ref. 16.
Note: Bold values signify the best performing methods.

values for noncortical structures for three methods: proposed, a patch-based,[10] and a random forest-based method.[12] It can be observed that the proposed method gives equal or marginally better performance than the patch-based method[10] on 13 out of 18 structures, while it is consistently better than random forest-based method[12] for all structures.

The FCN-based CNN method in Refs. 14 and 16 also report on this dataset for subcortical structure segmentation. A comparison of their reported DC values and ours is shown in Table 5. BrainSegNet appears to obtain a superior performance compared to both Refs. 14 and 16. It should be noted that these methods utilize postprocessing steps, such as MRF and random walker. A reason for the increased performance of our method, which has no postprocessing, can be attributed to the better use of contextual information in the proposed architecture.

#### 4.2.2 *LONI-LPBA40 dataset*

In this dataset,[31] largely "cortical" structures are labeled. BrainSegNet has a mean DC value of $0.824 \pm 0.040$ for this

**Table 6** Quantitative comparison on LPBA40 dataset. Reference 12 is based on random forest. DC values are listed for 54 structures.

| | Left brain | | Right brain | |
|---|---|---|---|---|
| | Ref. 12 | BrainSegNet | Ref. 12 | BrainSegNet |
| Superior frontal gyrus | $0.86 \pm 0.024$ | $\mathbf{0.89 \pm 0.023}$* | $0.86 \pm 0.020$ | $\mathbf{0.89 \pm 0.024}$* |
| Middle frontal gyrus | $0.85 \pm 0.029$ | $\mathbf{0.89 \pm 0.031}$* | $0.85 \pm 0.031$ | $\mathbf{0.88 \pm 0.026}$* |
| Inferior frontal gyrus | $0.79 \pm 0.046$ | $\mathbf{0.85 \pm 0.025}$* | $0.80 \pm 0.034$ | $\mathbf{0.85 \pm 0.036}$* |
| Precentral gyrus | $0.81 \pm 0.042$ | $\mathbf{0.86 \pm 0.040}$* | $0.82 \pm 0.039$ | $\mathbf{0.84 \pm 0.047}$* |
| Middle orbitofrontal gyrus | $0.75 \pm 0.069$ | $\mathbf{0.81 \pm 0.067}$* | $0.75 \pm 0.068$ | $\mathbf{0.80 \pm 0.056}$* |
| Lateral orbitofrontal gyrus | $0.69 \pm 0.096$ | $\mathbf{0.76 \pm 0.061}$* | $0.70 \pm 0.073$ | $\mathbf{0.72 \pm 0.082}$ |
| Gyrus rectus | $0.76 \pm 0.051$ | $\mathbf{0.79 \pm 0.068}$* | $0.75 \pm 0.051$ | $\mathbf{0.79 \pm 0.082}$ |
| Postcentral gyrus | $0.77 \pm 0.052$ | $\mathbf{0.83 \pm 0.051}$* | $0.78 \pm 0.071$ | $\mathbf{0.84 \pm 0.041}$* |
| Superior parietal gyrus | $0.80 \pm 0.040$ | $\mathbf{0.85 \pm 0.031}$* | $0.81 \pm 0.029$ | $\mathbf{0.84 \pm 0.029}$* |
| Supramarginal gyrus | $0.74 \pm 0.066$ | $\mathbf{0.79 \pm 0.054}$* | $0.75 \pm 0.073$ | $\mathbf{0.78 \pm 0.070}$* |
| Angular gyrus | $0.75 \pm 0.041$ | $\mathbf{0.76 \pm 0.070}$ | $0.74 \pm 0.070$ | $\mathbf{0.79 \pm 0.045}$* |
| Precuneus | $0.77 \pm 0.043$ | $\mathbf{0.79 \pm 0.064}$ | $0.77 \pm 0.039$ | $\mathbf{0.82 \pm 0.053}$* |
| Superior occipital gyrus | $0.69 \pm 0.075$ | $\mathbf{0.73 \pm 0.090}$* | $0.71 \pm 0.073$ | $\mathbf{0.72 \pm 0.081}$ |
| Middle occipital gyrus | $0.77 \pm 0.048$ | $\mathbf{0.79 \pm 0.064}$ | $0.78 \pm 0.050$ | $\mathbf{0.80 \pm 0.063}$ |
| Inferior occipital gyrus | $0.75 \pm 0.056$ | $\mathbf{0.81 \pm 0.056}$* | $0.76 \pm 0.057$ | $\mathbf{0.82 \pm 0.054}$* |
| Cuneus | $0.74 \pm 0.072$ | $\mathbf{0.83 \pm 0.067}$* | $0.75 \pm 0.067$ | $\mathbf{0.78 \pm 0.082}$ |
| Superior temporal gyrus | $0.84 \pm 0.027$ | $\mathbf{0.87 \pm 0.031}$* | $0.84 \pm 0.038$ | $\mathbf{0.88 \pm 0.033}$* |
| Middle temporal gyrus | $0.78 \pm 0.040$ | $\mathbf{0.81 \pm 0.045}$* | $0.76 \pm 0.046$ | $\mathbf{0.83 \pm 0.037}$* |
| Inferior temporal gyrus | $0.78 \pm 0.051$ | $\mathbf{0.83 \pm 0.047}$* | $0.76 \pm 0.048$ | $\mathbf{0.83 \pm 0.039}$* |
| Parahippocampal gyrus | $0.79 \pm 0.039$ | $\mathbf{0.82 \pm 0.053}$* | $0.79 \pm 0.036$ | $\mathbf{0.83 \pm 0.044}$* |
| Lingual gyrus | $0.80 \pm 0.054$ | $\mathbf{0.86 \pm 0.028}$* | $0.79 \pm 0.057$ | $\mathbf{0.85 \pm 0.041}$* |
| Fusiform gyrus | $0.80 \pm 0.051$ | $\mathbf{0.85 \pm 0.055}$* | $0.81 \pm 0.044$ | $\mathbf{0.86 \pm 0.044}$* |
| Insular cortex | $0.84 \pm 0.027$ | $\mathbf{0.88 \pm 0.033}$* | $0.86 \pm 0.020$ | $\mathbf{0.87 \pm 0.038}$ |
| Cingulate gyrus | $0.77 \pm 0.065$ | $\mathbf{0.79 \pm 0.042}$ | $0.79 \pm 0.044$ | $\mathbf{0.80 \pm 0.051}$ |
| Caudate | $0.81 \pm 0.046$ | $\mathbf{0.84 \pm 0.060}$* | $0.81 \pm 0.064$ | $\mathbf{0.82 \pm 0.076}$ |
| Putamen | $0.82 \pm 0.028$ | $\mathbf{0.83 \pm 0.039}$ | $0.81 \pm 0.028$ | $\mathbf{0.85 \pm 0.043}$* |
| Hippocampus | $0.81 \pm 0.026$ | $\mathbf{0.83 \pm 0.036}$* | $0.81 \pm 0.048$ | $\mathbf{0.83 \pm 0.038}$* |
| Overall | $0.78 \pm 0.048$ | $\mathbf{0.82 \pm 0.039}$* | $0.79 \pm 0.048$ | $\mathbf{0.82 \pm 0.041}$* |

*The statistically significance difference ($p < 0.05$) between BrainSegNet and Ref. 12.
Note: Bold values signify the best performing methods.

**Table 7** Quantitative comparison on two variants (Hammers67n20 and Hammers83n30) of IXI database. DC values are given for subcortical structures that are common for both the datasets.

| | | BrainSegNet | |
|---|---|---|---|
| | Ref. 24 | Hammers67n20 | Hammers83n30 |
| Hippocampus | **0.85** | $0.84 \pm 0.038$ | $0.83 \pm 0.035$ |
| Amygdala | **0.82** | $0.81 \pm 0.089$ | $0.81 \pm 0.070$ |
| Caudate | **0.90** | $\mathbf{0.90 \pm 0.029}$ | $0.88 \pm 0.026$ |
| Nucleus accumbens | **0.71** | $0.70 \pm 0.112$ | $0.68 \pm 0.105$ |
| Putamen | **0.89** | $\mathbf{0.89 \pm 0.031}$ | $\mathbf{0.89 \pm 0.039}$ |
| Thalamus | 0.90 | $\mathbf{0.91 \pm 0.015}$ | $0.90 \pm 0.021$ |
| Pallidum | 0.80 | $\mathbf{0.81 \pm 0.076}$ | $0.80 \pm 0.066$ |

Note: Bold values signify the best performing methods.

set, which is higher than $0.814 \pm 0.023$ ($p < 0.01$) reported by the current state-of-the-art patch-based technique.[24] The random forest-based method[12] also uses a classification approach, such as the proposed method. The DC values for 54 structures in LPBA40 dataset are listed in Table 6 for comparison. Based on these values, BrainSegNet appears to perform consistently better.

#### 4.2.3 *IXI datasets*

IXI database consists of two datasets with 67 and 83 labeled structures as mentioned in Sec. 3.

Hammers67n20 Dataset:[2] BrainSegNet obtains a mean DC of $0.840 \pm 0.013$ on this dataset, which is comparable to $0.836 \pm 0.009$ ($p = 0.35$) reported in Ref. 2 with nonrigid registration. Reference 2 also reports a lower figure of $0.754 \pm 0.016$ ($p < 0.01$) with affine registration. It should be noted that though Ref. 2 reports on 30 MRI scans, only 20 scans are available for public access.

Hammers83n30 Dataset:[2] BrainSegNet has a mean DC of $0.808 \pm 0.014$ on this dataset, which is better than that obtained by different nonrigid registration-based methods reported in Ref. 37, such as demons (0.785, $p < 0.01$), PCA (0.754, $p < 0.01$), HAMMER (0.789, $p < 0.01$), and ISA (0.801, $p < 0.01$).

Table 7 lists the DC values obtained with the proposed method (BrainSegNet) and those reported in Ref. 24 for subcortical structures common to both of the above two datasets. It is desirable for a method to perform equally well on the same structures from both datasets. This is true for our method. It can also be observed that BrainSegNet has equal or marginally higher DC than Ref. 24 for four out of seven structures while it is marginally less for the rest.

Finally, the generalizability of the obtained results with BrainSegNet was also assessed. Training was done with the MICCAI-2012 dataset and testing was done on the IXI datasets (Hammers67n20 and Hammers83n30). As subcortical structures are common between these three datasets, the performance is compared for only these structures in Table 8. The tabulated results indicate only marginal degradation with cross training, which implies that the proposed method is robust to change in dataset. This eliminates the need for retraining for a new dataset.

## 5 Discussion and Conclusion

Traditional approaches, such as MALP, for the brain structure segmentation rely on nonrigid registration to label a new test volume. Various machine learning-based algorithms, such as random forest[11] and CNN,[13–16] treat structure segmentation as a classification task. These existing CNN-based methods, however, have the following drawbacks: multiple passes over a test image to obtain final labels and dependency on initial model;[13] need for postprocessing (such as MRF,[14] Hough voting,[15] or random walker[16]), which prohibits an end-to-end trainable CNN-based solution.

We presented an end-to-end trainable CNN-based solution (BrainSegNet) for brain structure segmentation by employing 2-D/3-D patches of varying size as input. The experimental results of the variants of the proposed method (see Table 3) demonstrated that both context (branches 2 and 4) and appearance (branches 1 and 3) are important for labeling, with wider context boosting the DC values of cortical more than noncortical structures.

**Table 8** Quantitative comparison of two variants (self-training and cross-training) on IXI databases. DC values are given for subcortical structures that are common for both the datasets.

| | Self-training (same dataset) | | Cross-training (MICCAI-2012 dataset) | |
|---|---|---|---|---|
| | Hammers67n20 | Hammers83n30 | Hammers67n20 | Hammers83n30 |
| Hippocampus | $0.84 \pm 0.038$ | $0.83 \pm 0.035$ | $0.82 \pm 0.042$ | $0.82 \pm 0.039$ |
| Amygdala | $0.81 \pm 0.089$ | $0.81 \pm 0.070$ | $0.80 \pm 0.096$ | $0.80 \pm 0.078$ |
| Caudate | $0.90 \pm 0.029$ | $0.88 \pm 0.026$ | $0.89 \pm 0.022$ | $0.88 \pm 0.029$ |
| Nucleus accumbens | $0.70 \pm 0.112$ | $0.68 \pm 0.105$ | $0.69 \pm 0.119$ | $0.69 \pm 0.101$ |
| Putamen | $0.89 \pm 0.031$ | $0.89 \pm 0.039$ | $0.88 \pm 0.043$ | $0.89 \pm 0.043$ |
| Thalamus | $0.91 \pm 0.015$ | $0.90 \pm 0.021$ | $0.90 \pm 0.017$ | $0.89 \pm 0.023$ |
| Pallidum | $0.81 \pm 0.076$ | $0.80 \pm 0.066$ | $0.80 \pm 0.081$ | $0.80 \pm 0.073$ |

The results on five different publicly available datasets and especially the IXI datasets (with 67 and 84 labeled structures) indicate that DC tends to degrade with an increase in the number of labeled structures in a dataset. This suggests that the source of degradation could potentially be either a class-imbalance problem, which arises in an MAS problem with a large number of classes, or the chosen evaluation metric. Hence, we computed the "accuracy" of segmentation for the two IXI datasets with 67 and 83 structures. It was found that while the mean accuracy was 0.868 for both datasets, the mean DC values were different. This can be attributed to the fact that the loss function optimizes accuracy and not DC. Thus, the degradation appears to be due to the evaluation metric (which was chosen based on its popularity in literature) and not the class imbalance. In the future, it would be of interest to explore a loss function which optimizes DC to confirm the above observation.

It was also observed that labeling error was higher in the brain-skull boundary region. Hence, a brain mask was generated using the "GroundTruth" mask. With this modification, the mean DC was found to increase by ∼1% for all the datasets (new values for MICCAI: 0.7542, IBSR: 0.853, LONI-LPBA40: 0.834, Hammers67n20: 0.849, and Hammers83n30: 0.815). Hence, skull stripping does play a role in performance.

The proposed BrainSegNet adopts a classification-based approach for brain structure segmentation. A comparative analysis done against other classification-based approaches, patch-based, and registration-based methods on four datasets showed that BrainSegNet has comparable or marginally better performance and at a reduced computational time (see Table 2). Experimental results, in Sec. 4.2.3, showed it to be robust to changes in datasets obviating the need to retrain when the structures of interest are common across the datasets. This coupled with the fact that a nonrigid registration step is not required should make the proposed solution attractive for many applications where computational time plays a critical role.

## Disclosures

Raghav Mehta, Aabhas Majumdar, and Jayanthi Sivaswamy have no conflicts of interest, financial or otherwise.

## References


1. T. R. Langerak et al., "Multiatlas-based segmentation with preregistration atlas selection," *Med. Phys.* **40**(9), 091701 (2013).
2. R. A. Heckemann et al., "Automatic anatomical brain MRI segmentation combining label propagation and decision fusion," *NeuroImage* **33**(1), 115–126 (2006).
3. H. Wang, B. Avants, and P. Yushkevich, "A combined joint label fusion and corrective learning approach," in *MICCAI Workshop on Multi-Atlas Labeling* (2012).
4. J. E. Iglesias and M. R. Sabuncu, "Multi-atlas segmentation of biomedical images: a survey," *Med. Image Anal.* **24**(1), 205–219 (2015).
5. B. Landman and S. Warfield, "MICCAI 2012 workshop on multi-atlas labeling," in *Medical Image Computing and Computer Assisted Intervention Conf. 2012: MICCAI 2012 Grand Challenge and Workshop on Multi-Atlas Labeling Challenge Results* (2012).
6. A. Asman and B. Landman, "Multi-atlas segmentation using non-local staple," in *MICCAI Workshop on Multi-Atlas Labeling* (2012).
7. A. K. H. Duc et al., "Using manifold learning for atlas selection in multi-atlas segmentation," *PLoS One* **8**(8), e70059 (2013).
8. P. Coupé et al., "Patch-based segmentation using expert priors: application to hippocampus and ventricle segmentation," *NeuroImage* **54**(2), 940–954 (2011).
9. D. Zhang et al., "Sparse patch-based label fusion for multi-atlas segmentation," in *Int. Workshop on Multimodal Brain Image Analysis*, pp. 94–102, Springer (2012).
10. F. Rousseau, P. A. Habas, and C. Studholme, "A supervised patch-based approach for human brain labeling," *IEEE Trans. Med. Imaging* **30**(10), 1852–1862 (2011).
11. D. Zikic, B. Glocker, and A. Criminisi, "Atlas encoding by randomized forests for efficient label propagation," in *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention*, pp. 66–73, Springer (2013).
12. L. Zhang et al., "Automatic labeling of MR brain images by hierarchical learning of atlas forests," *Med. Phys.* **43**(3), 1175–1186 (2016).
13. A. de Brebisson and G. Montana, "Deep neural networks for anatomical brain segmentation," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition Workshops*, pp. 20–28 (2015).
14. M. Shakeri et al., "Sub-cortical brain structure segmentation using F-CNN's," in *IEEE 13th Int. Symp. on Biomedical Imaging (ISB)*, pp. 269–272 (2016).
15. F. Milletari et al., "Hough-CNN: deep learning for segmentation of deep brain regions in MRI and ultrasound," arXiv preprint arXiv:1601.07014 (2016).
16. S. Bao and A. C. Chung, "Multi-scale structured CNN with label consistency for brain MR image segmentation," *Comput. Methods Biomech. Biomed. Eng.: Imaging Visualization*, **4**, 1–5 (2016).
17. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, pp. 1097–1105 (2012).
18. J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 3431–3440 (2015).
19. H. R. Roth et al., "Deep convolutional networks for pancreas segmentation in CT imaging," *Proc. SPIE* **9413**, 94131G (2015).
20. O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, Springer (2015).
21. W. Zhang et al., "Deep convolutional neural networks for multi-modality isointense infant brain image segmentation," *NeuroImage* **108**, 214–224 (2015).
22. M. F. Stollenga et al., "Parallel multi-dimensional LSTM, with application to fast biomedical volumetric image segmentation," in *Advances in Neural Information Processing Systems*, pp. 2998–3006 (2015).
23. P. Moeskops et al., "Automatic segmentation of MR brain images with a convolutional neural network," *IEEE Trans. Med. Imaging* **35**(5), 1252–1261 (2016).
24. G. Wu et al., "Hierarchical multi-atlas label fusion with multi-scale feature representation and label-specific patch partition," *NeuroImage* **106**, 34–46 (2015).
25. F. Rousseau, P. A. Habas, and C. Studholme, "A supervised patch-based approach for human brain labeling," *IEEE Trans. Med. Imaging* **30**(10), 1852–1862 (2011).
26. N. Srivastava et al., "Dropout: a simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.* **15**(1), 1929–1958 (2014).
27. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Int. Conf. on Learning Representations (ICLR)* (2014).
28. K. Kamnitsas et al., "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," *Med. Image Anal.*, **36**, 61–78 (2017).
29. https://masi.vuse.vanderbilt.edu/workshop2012.
30. D. S. Marcus et al., "Open access series of imaging studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults," *J. Cognit. Neurosci.* **19**(9), 1498–1507 (2007).
31. D. W. Shattuck et al., "Construction of a 3D probabilistic atlas of human cortical structures," *NeuroImage* **39**(3), 1064–1080 (2008).
32. http://loni.usc.edu/atlases/Atlas_Methods.php?atlas_id=12.
33. http://brain-development.org/brain-atlases/.
34. N. J. Tustison et al., "N4ITK: improved N3 bias correction," *IEEE Trans. Med. Imaging* **29**(6), 1310–1320 (2010).
35. S. M. Smith, "Fast robust automated brain extraction," *Hum. Brain Mapp.* **17**(3), 143–155 (2002).

36. http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/BET.
37. G. Wu et al., "Unsupervised deep feature learning for deformable registration of MR brain images," in *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention*, pp. 649–656, Springer (2013).

**Raghav Mehta** received the BE degree in electronics engineering from the Birla Vishvakarma Mahavidyalay Engineering College, Vallabh Vidhyanagar, Gujarat, India, in 2014. He is currently working toward the MS (by research) degree in electronics and communication engineering at the Center for Visual Information (CVIT), International Institute of Information Technology (IIIT), Hyderabad, India. His research interests include image processing, neuro image analysis, and machine learning.

**Aabhas Majumdar** is currently working toward dual degree (Btech and MS by research) in computer science at CVIT, IIIT, Hyderabad, India, under the supervision of Jayanthi Sivaswamy. His research interests are image registration and general-purpose computing on GPUs.

**Jayanthi Sivaswamy** received her PhD in electrical engineering from Syracuse University, New York, USA. Since 2001, she has been with IIIT, Hyderabad, India. Prior to that, she was at the University of Auckland, Auckland, New Zealand. Her research interests are CAD algorithm development and neuroimage analysis and image reconstruction.