



Fully Automatic Segmentation for Ischemic Stroke Using CT Perfusion Maps

Vikas Kumar Anand, Mahendra Khened, Varghese Alex,
and Ganapathy Krishnamurthi^(✉)

Indian Institute of Technology Madras, Chennai, India
gankrish@iitm.ac.in

Abstract. We propose an algorithm for automatic segmentation of ischemic lesion using CT perfusion maps. Our method is based on encoder-decoder fully convolutional neural network approach. The pre-processing step involves skull stripping and standardization of perfusion maps and extraction of slices with lesions as the training data. These CT perfusion maps are used to train the proposed network for automatic segmentation of stroke lesions. The network is trained by minimizing the weighted combination of cross entropy and dice losses. Our algorithm achieves 0.43, 0.53 and 0.45 Dice, precision, and recall respectively on challenge test data set.

Keywords: Deep learning · CNN · CT perfusion · Ischemic stroke

1 Introduction and Related Work

Ischemic stroke is one of the most frequent cerebrovascular disease and vital cause of death and disability in the world. Defining location and extent of irreversibly damaged brain tissue is a critical part of the decision-making process in acute stroke. When compared to normal tissues, the infarcted tissues in the volume differ on the basis of diffusion & perfusion, hence making diffusion and perfusion MR imaging the gold standard for ischemic lesions. Recently, due to high speed, availability and lack of contradiction, Computed Tomography (CT) has been used to evaluate the order of treatment for ischemic stroke. Though the contrast of ischemic stroke lesions are poor in brain CT images, CT perfusion imaging could be exploited to identify & locate the ischemic stroke lesions. Qualitative images derived from CT perfusion imaging comprises of Cerebral Blood Flow (CBF), Cerebral Blood Volume (CBV), time to peak of the residue function (Tmax) and Mean Transit Time (MTT). CBV is defined as the volume of blood in a given amount of brain tissue. CBF is defined as CBV per unit time. The ratio of CBV and CBF is expressed as MTT.

The work presented in the paper is an attempt to develop an algorithm for fully automatic segmentation for ischemic stroke using CT perfusion maps. Several perfusion maps (CBV, CBF, MTT, Tmax) were used as inputs to the algorithm whereas ground truth used was generated from MRI Diffusion Weighted

Imaging (DWI) of ischemic stroke lesion. On DWI maps, infarcted brain tissues are portrayed as hyper-intense regions.

Convolutional neural networks (CNNs) [7] have been applied to wide variety of pattern recognition tasks, most common ones are image classification [2, 6, 12] and semantic segmentation using fully convolutional networks (FCN) [8]. CNNs have also been applied to medical image segmentation and classification [1, 9]. In this paper, we propose a CNN based architecture for segmentation of the stroke lesion from CT perfusion maps and CT images of the brain. Our network’s connectivity pattern was inspired from DenseNets [3].

2 Materials and Methods

2.1 Data Pre-processing Pipeline

Skull Stripping and Brain Mask Generation. Skull stripping is an important pre-processing step in brain image analysis. In this process, non brain tissues like skull, eyes, scalp, dura were removed. This process helped in enhancing the segmentation accuracy and lowering the execution time required by segmentation algorithms. Brain extraction tool (BET) [13] was utilized for removal of skull and another non brain tissues. This was followed by Hounsfield windowing (0–100) of brain tissues so as to improve the contrast.

A brain mask was generated from skull stripped CT images of brain by means of simple thresholding. The brain mask was used to remove non-brain regions from CBF, CBV, MTT and Tmax maps. Different steps in the data pre-processing pipeline are illustrated in Fig. 1.

Data Normalization. Volumes were normalized to have zero mean and unit variance using Eq. 1.

$$X_{norm} = \frac{X - \mu}{\sigma} \quad (1)$$

where X is the data, μ and σ are global mean and global standard deviation associated with X . These were calculated from entire training dataset.

2.2 Proposed Network

We utilize an encoder-decoder architecture for the task of segmenting ischemic lesion from the CT perfusion maps. The encoder in the network comprise of dense connectivity pattern & was inspired from the DenseNet-121 architecture. The decoder was composed of the bi-linear up-sampling module and convolutional layers. Features learnt in the down-sampling path were concatenated with the features learnt in up-sampling path using long skip connection. The architecture of the network is illustrated in Fig. 2.

The input to the network comprised of 5 channels for feeding CT image and corresponding 4 different CT perfusion maps. The first convolutional layer composed of 64 different 7×7 kernels. The resultant feature maps were then passed

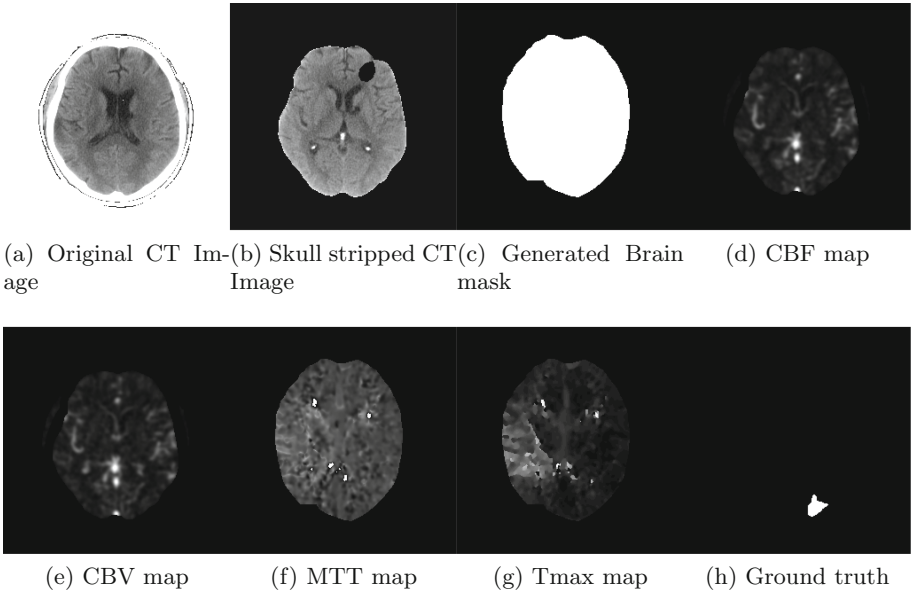


Fig. 1. Data pre-processing steps involved (a) CT image after Hounsfield windowing, (b) Skull stripped image, (c) Mask generated using skull stripped image, (d) Cerebral blood flow information, (e) Cerebral blood volume information, (f) MTT, (g) Tmax and (h) Ground-truth image

to a batch normalization [4] and a non linearity layer (ReLU) [11]. This was followed by a max pooling layer with the kernel size and stride set to 3×3 and 2 respectively. Max-pooling aided in reducing the spatial dimension of the generated features. The features were passed through a series of dense block and

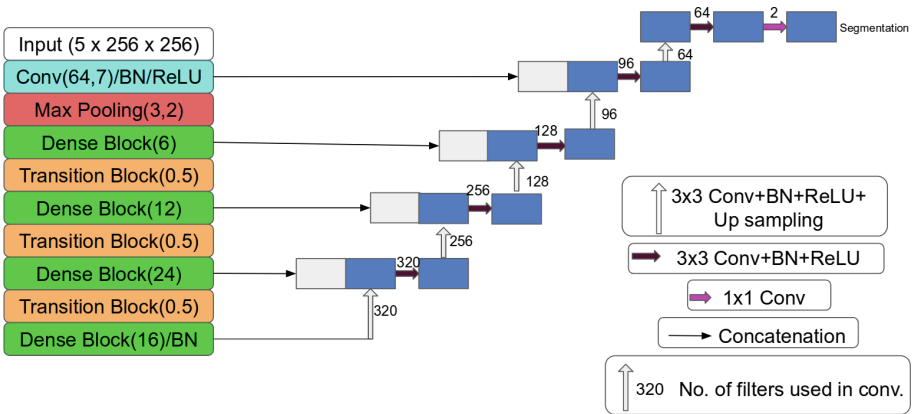


Fig. 2. Proposed network

transition layers. A dense block was composed of convolutional layers, wherein each convolutional layer received inputs from all the preceding layers in the network, Fig. 3. In a dense block, batch normalization and a non-linearity layer preceded a convolutional layer in the network, while a 2-D dropout layer ($p = 0.5$) succeeded the convolutional layer, Fig. 4. In a dense block, the number of convolutional layers was an hyper-parameter. In this work, the network composed of 4 dense blocks and had 6, 12, 24, and 16 convolutional layers respectively. Transition Down block was composed of $[1 \times 1]$ convolutional layer followed by dropout with $p = 0.5$ and $[2 \times 2]$ max pooling layer and aided in reducing the spatial dimension of the features learnt by the network Fig. 5.

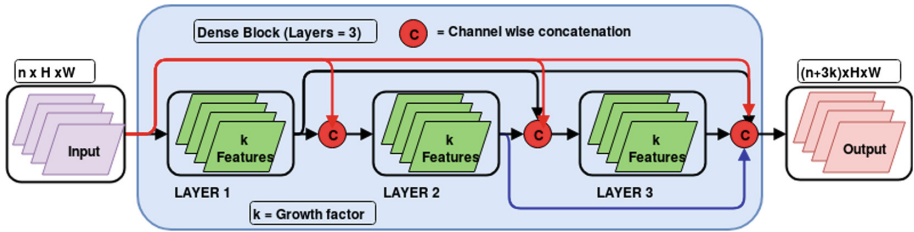


Fig. 3. Dense block with 3 layers

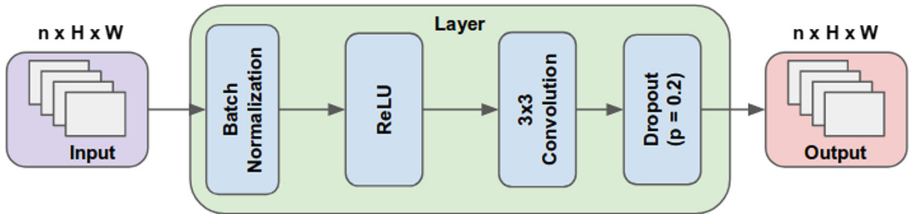


Fig. 4. Layer of dense block

2.3 Loss Function

Lesions are represented by a minuscule proportion of voxels in a medical volume by thereby leading to class imbalance. This issue was circumvented by training the network to minimize a hybrid loss function. The hybrid cost function comprised of weighted cross entropy & dice loss [10].

The dice co-efficient is an overlap metric used for assessing the quality of segmentation maps. The dice coefficient between two binary volumes can be written as:

$$DICE = \frac{2 \sum_i^N p_i g_i}{\sum_i^N p_i^2 + \sum_i^N g_i^2} \tag{2}$$

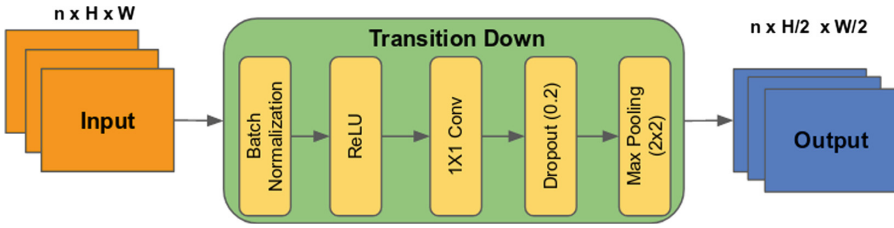


Fig. 5. Transition block

where the sums run over the N voxels, of the predicted binary segmentation volume $p_i \in P$ and the ground truth binary volume $g_i \in G$.

The parameters of the network was optimized so as to minimize the *total_loss*, Eq. (3).

$$total_loss = \lambda(cross_entropy_loss) + \gamma(dice_loss_bg) + \delta(dice_loss_fg) \quad (3)$$

where λ , δ and γ are empirically assigned weights to individual losses, fg and bg represent foreground voxels which corresponded to lesion regions and background voxels which corresponded to non-lesion regions respectively. In this work we set $\gamma = 0.25$, $\delta = 0.25$ and $\lambda = 0.50$.

Training. The proposed model was trained on a batch size of 6 for 90 epochs using ADAM [5] as the optimizer. The provided dataset was split into training and validation subsets in the ratio of 7:3. The weights of the encoding layers of our network were initialized using DenseNet-121 architecture pre-trained on ImageNet dataset. Since DenseNet-121 architecture took 3 channel input, the weights of the first layer of our network were randomly initialized. During training of our network, the model was saved on every epoch and the best model selection criteria was based on the model which gave the highest dice score on the validation set.

3 Experimental Setup and Result

3.1 Data Sets and Evaluation Criteria

Imaging data from acute stroke patients in two centers who presented within 8 h of stroke onset and underwent an MRI DWI within 3 h after CT perfusion were included. To assess cerebral perfusion, a contrast agent (CA) was administered to the patient and its temporal change was captured in dynamic scans acquired 1–2 s apart. There were 63 patients in the training data set. Each patient data comprised of brain CT image, CBV map, CBF map, MTT map, Tmax map with ground truth that was generated on DWI map. To access our segmentation algorithm's performance, the organizers employed Dice, precision, recall, Hausdroff distance, Average distance and Absolute volume difference.

3.2 Experimental Results

See Fig. 6 and Table 1.

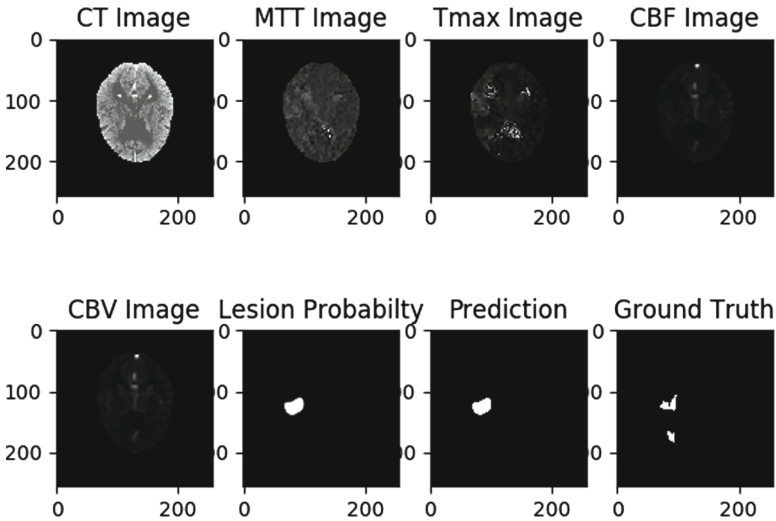


Fig. 6. The figure illustrates our model’s prediction on held-out test set.

Table 1. The table provides the list of evaluation metrics used for evaluating our segmentation on the challenge test dataset ($n = 62$). HD - Hausdorff Distance, AD - Average Distance, AVD - Absolute Volume Difference. The values provided are mean (standard deviation).

Dice	Precision	Recall
0.43 (0.30)	0.53 (0.38)	0.45 (0.33)
HD	AD	AVD
16129052.53 (36779843.53)	16791858.37 (36853432.68)	13.13 (14.82)

3.3 Conclusion

We developed an automatic segmentation algorithm for ischemic stroke lesions using CT perfusion maps. CT imaging has several advantages such as less scan time, low cost and ease of availability over MRI. The analysis of CT perfusion maps have shown to be useful for early treatment planning on the onset of stroke.

References

1. Ciresan, D., Giusti, A., Gambardella, L.M., Schmidhuber, J.: Deep neural networks segment neuronal membranes in electron microscopy images. In: *Advances in Neural Information Processing Systems*, pp. 2843–2851 (2012)
2. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778 (2016)
3. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: *CVPR*, vol. 1, p. 3 (2017)
4. Ioffe, S., Szegedy, C.: Batch normalization: accelerating deep network training by reducing internal covariate shift. *arXiv preprint [arXiv:1502.03167](https://arxiv.org/abs/1502.03167)* (2015)
5. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. *arXiv preprint [arXiv:1412.6980](https://arxiv.org/abs/1412.6980)* (2014)
6. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: *Advances in Neural Information Processing Systems*, pp. 1097–1105 (2012)
7. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. *Proc. IEEE* **86**(11), 2278–2324 (1998)
8. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440 (2015)
9. Menze, B.H., et al.: The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans. Med. Imaging* **34**(10), 1993 (2015)
10. Milletari, F., Navab, N., Ahmadi, S.A.: V-NET: fully convolutional neural networks for volumetric medical image segmentation. In: *2016 Fourth International Conference on 3D Vision (3DV)*, pp. 565–571. IEEE (2016)
11. Nair, V., Hinton, G.E.: Rectified linear units improve restricted Boltzmann machines. In: *Proceedings of the 27th International Conference on Machine Learning (ICML 2010)*, pp. 807–814 (2010)
12. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. *arXiv preprint [arXiv:1409.1556](https://arxiv.org/abs/1409.1556)* (2014)
13. Smith, S.M.: Fast robust automated brain extraction. *Hum. Brain Mapp.* **17**(3), 143–155 (2002)