



# Multi-scale Deep Convolutional Neural Network for Stroke Lesions Segmentation on CT Images

Liangliang Liu<sup>1</sup>, Shuai Yang<sup>2</sup>, Li Meng<sup>2</sup>, Min Li<sup>1</sup>, and Jianxin Wang<sup>1</sup>(✉)

<sup>1</sup> School of Information Science and Engineering, Central South University, Changsha 410083, People's Republic of China

liuliang.double@csu.edu.cn, {limin, jxwang}@mail.csu.edu.cn

<sup>2</sup> Department of Radiology, Xiangya Hospital, Central South University, Changsha 410008, People's Republic of China

**Abstract.** Ischemic stroke is the top cerebral vascular disease leading to disability and death worldwide. Accurate and automatic segmentation of lesions of stroke can assist diagnosis and treatment planning. However, manual segmentation is a time-consuming and subjective for neurologists. In this study, we propose a novel deep convolutional neural network, which is developed for the segmentation of stroke lesions from CT perfusion images. The main structure of network bases on U-shape. We embed the dense blocks into U-shape network, which can alleviate the over-fitting problem. In order to acquire more receptive fields, we use multi-kernel to divide the network into two paths, and use the dropout regularization method to achieve effective feature mapping. In addition, we use multi-scale features to obtain more spatial features, which will help improve segmentation performance. In the post-processing stage of soft segmentation, we use image median filtering to eliminate the specific noises and make the segmentation edge smoother. We evaluate our method in Ischemic Stroke Lesion Segmentations Challenge (ISLES) 2018. The results of our approach on the testing data places high ranking.

**Keywords:** Stroke · CT perfusion images · Dropout · Multi-scale · U-shape network

## 1 Introduction

The stroke, one of the leading causes of death and disability worldwide, is triggered by an obstruction in the cerebrovascular system preventing the blood to reach the brain regions supplied by the blocked blood vessel directly. Ischemic stroke is the commonest subtype of stroke, which is a disease with sudden onset and high mortality. It prevents blood flow in small vessels. When the blood flow interruption is too long, cell will undergo necrosis and irreversibly injured infarct core is formed [7]. Defining location and extend of the infarct core is a critical

part of the decision making process in acute stroke. In clinical diagnosis, CT image is a speed, availability, and lack of contraindications manner to triage stroke patients. If we can locate the location and size of the lesion quickly, it is the key to save some viable brain tissue [24]. In traditional medical diagnosis, the lesion tissue is accomplished by manual segmentation on medical images. However, manual delineation of stroke lesions a time-consuming and very tedious task [8]. Automatic and accurate quantification of stroke lesions is an important metric for planning treatment strategies, monitoring disease progression.

Over the past decades, Unsupervised methods and shallow machine learning methods are traditional methods of image analysis, such as: multi-modal generative based mixture-model [1], image cross-saliency approach [3], spatial decision forests approach [5] and multi-atlas segmentation method [19], and so on, those methods had been successful. However, there are also some limitations in these methods. For example, some of those methods are designed specifically require and heavily dependent on handcrafted lesions segmentation [11, 12] or improve the accuracy of segmentation depend on multi-atlas label [23].

Recent years, deep convolution neural networks (DCNNs) are one of the most competitive approach used for medical image semantic segmentation. The DCNN models are capable of learning features from raw images and extracting context information. The feature sets filtered by DCNN often outperform pre-defined and hand-crafted feature sets. For example, Ronneberger et al. proposed a novel U-net model based on DCNN architecture [25]. U-net combined the down-sampling layers and up-sampling layers with skip connections, this architecture can reuse the context information of the down-sampling layers and greatly improve the performance of the segmentation. Long et al. proposed a novel framework to automatically segment stroke lesions. This framework consists of two deep convolutional neural networks, and it achieved state-of-the-art performance on an acute ischemic stroke dataset [21]. Zhang et al. used a custom DCNN to automatic segmentation acute ischemic stroke from DWI modality, in the network, they used dense connectivity to relieve the problems of deep network, and the network outperforms other state-of-the-art methods by a large margin [28]. Li et al. developed an automatic intervertebral discs (IVDs) segmentation method based on fully convolution networks [20], they used multi-scale and feature dropout learning technology to segment region of interest (ROI) from multi-modality MRI images, this method achieved the 1st place in the MICCAI challenge in 2016. Others methods based on DCNN which are applied in medical images of various diseases, such as: stroke image segmentation [22], brain tumor image segmentation [17], WMH segmentation [9], and optic disc segmentation [4], and so on. Most these methods are based on magnetic resonance imaging (MRI). Especially, the segmentation methods of stroke lesions is seldom used in CT images.

In this paper, we propose a novel multi-scale features deep convolution neural network (MS-DCNN) for stroke lesions segmentation on CT images. The whole neural network consists of a series of convolution layers, dense blocks [13], transition blocks and upsampling blocks. We use the dropout regularization method

to alleviate neural network from over-fitting. We use random rotated and distortion to increase the number of training samples. He network with the main contributions as follows:

1. We propose an end-to-end deep convolution neural network base on two symmetrical U-shape networks [25], and embedded dense blocks into the U-shape [13]. This strategy can improve the information on the sampling and improve the feature reuse.
2. We use the dropout regularization method in dense blocks and transition blocks. It's a simple method to prevent neural network from over-fitting and improve the neural network efficiency. Proper use of dropout can help improve the accuracy of segmentation.
3. We employ dual parallel kernel pathways in our framework to process input CT images. This design can help extract the image features fully, and finally combine the two pathways before output, it helps to improve the performance of the segmentation [27]. We evaluate our method on ISLES 2018 challenge.

## 2 Material and Method

### 2.1 Data

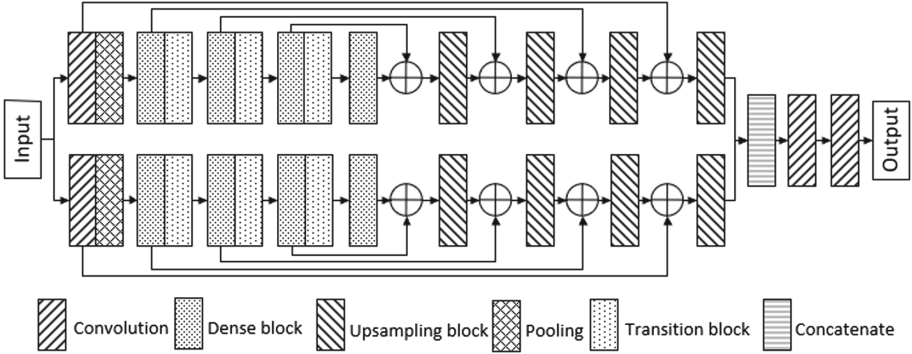
Ischemic Stroke Lesion Segmentations Challenge (ISLES) 2018 offers a platform for participants to compare their methods directly and fair. ISLES 2018 challenge offers 103 stroke patients, which is based on acute CT perfusion data. Each patient has 5 CT sequences (CBV, CBF, MTT, TMAX, CTP). Imaging data from acute stroke patients in two centers who presented within 8 h of stroke onset and underwent an Magnetic Resonance Imaging (MRI) DWI within 3 h after CTP were included. The challenge's training data set consists of 63 patients, some patient cases have two slabs to cover the stroke lesion, finally, we got 94 samples in the training dataset. The testing dataset consists of 40 patients. Some patient cases have two slabs to cover the stroke lesion. We got 62 testing samples. In this challenges, the training data set and the ground truth are opened to all participants. The testing data set only open the CT images which is to be predicted, without the ground truth is distributed on the challenge web pages. Participants should submitted their final segmentation results to the organizers, who scored the segmentation results.

### 2.2 MS-DCNN

A traditional image-processing CNN is composed of one input layer, many convolution layers and one output layer. Features are transmitted by single line between layers, which leads to inadequate extraction features. We propose the MK-DCNN framework is based on the U-net architecture [25] and we embed dense structure as a block into the U-shape framework [13], both two methods include jump layer which can help to improve feature reuse.

Figure 1 illustrates the pipeline of our proposed segmentation network. Our network is based on two symmetric U-shape structures, and we use dense block

to implement the down-sampling operation in contracting path of U-shape, and after completes up-sampling operation in expansive path, we concatenate two symmetric networks and output the predicted result. We use multi-scale features strategy to enhance the feature extraction sufficiently. In the first layer, we use dual parallel kernels in two symmetric pathway to extract different features. To handle the problem of over-fitting of DCNNs, we not only use dense block to resist over-fitting, but also use dropout regularization method to alleviate over-fitting and improve the efficiency of neural network.



**Fig. 1.** Architecture of the MS-DCNN.

As shown in Fig. 1. Our network consists of 3 separate convolution layers, 1 pooling layer, 4 dense blocks, 3 transition blocks and 4 up-sampling blocks. We extend the deep of DenseNet-121 to 123 layers in dense blocks. Each dense block contains several micro-dense units, each dense unit is composed of a batch normalization (BN) [16] layer, a rectified linear units (ReLU) [6] layer and a convolution (Conv) layer, the concatenation operation is required before result output. A n-layer dense block consists of a dense unit or several continuous dense units. Figures 2 and 3 illustrate the basic implementation of a dense unit and a n-layer dense block, respectively. In dense block, each dense unit is regarded as one layer, all layers inside the block are directly connected. The transition block consists of a BN layer, a ReLU layer and an average pooling layer [18]. We embed the dropout regularization method into the both dense block and transition block. The up-sampling block consists of a concatenate layer, a BN layer, a ReLU layer and a Conv layer, we use bilinear interpolation technology to realize image zooming. Then, we concatenate the two un-sampling results which come from different paths. Finally, after two convolution layers, we use the sigmoid function to complete segmentation task and output the final lesion information.

In our network, we only use 4 CT modality sequences (CBV, CBF, TMAX and MTT). According to the clinical prior knowledge, 4 modalities play different roles in stroke diagnosis, we divide the 4 modalities into two groups

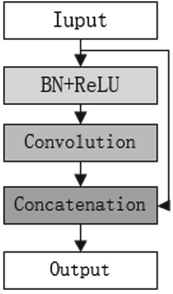


Fig. 2. A dense unit.

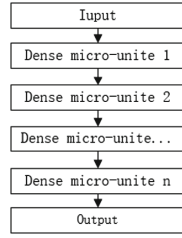


Fig. 3. Architecture of the n-layer dense block.

(TMAX+CBF+CBV and MTT). First, We set different dropout rates for the two groups in our network. Then, we concatenate the 2 unsampling results which come from 2 pathways of MS-DCNN. Finally, after two convolution operations, we use the sigmoid function to complete segmentation task and out the final lesion information.

### 2.3 Dropout Regularization Method for Effective Learning

The regularization is a popular method to prevent over-fitting and filter important features. It is a very important and effective technology to reduce generalization error in machine learning. Regularization can automatically weak the unimportant feature variables. Dropout is one of a general and concise regularization methods which performs well in many tasks [10,26]. In our study, we use dropout to reduce redundant features produced by multi-scale method and to alleviate the problem of duplicate feature acquisition from the same area of the image. We use dropout regularization method in dense block and transition block. The application of dropout on a generic  $i$ -th neuron in the  $n$  layer is shown below:

$$Q_i = x_i a\left(\sum_{k=1}^{d_i} w_k x_k + b_k\right) (0 \leq i \leq h), \tag{1}$$

where  $Q_i$  is the retained probability of the  $i$ -th neuron,  $x_i$  is the  $i$ -th neuron,  $a()$  is an activation function,  $k \in [1, i]$  is unit number,  $w_k$  and  $b_k$  are the  $k$ -th unit weight and bias.  $d$  denotes dimensional,  $x_{d_i}$  denotes  $x_i$  is a Bernoulli variables with  $d$  dimensional.  $\sum_{k=1}^{d_i} w_k x_k$  is the sum of the product of all neurons weights  $w_k$  and  $x_k$  before  $i$ -th neuron.

In our network, we need to dropout a set of neurons of a layer. Let the  $j$ -th layer has  $n$  neurons, in a cycle, the neural network can be regarded as the integration  $n$  times of Bernoulli's experiments, and the probability of each neuron being retained is  $q$  and the dropout probability is  $p$ . Thus, the number of neurons retained in layer  $j$ -th is as follows:

$$Y = \sum_{i=1}^{d_j} x_i, \quad (2)$$

where  $x_i$  is a retained neuron (a Bernoulli random variable). In the  $n$  experiments, the probability of retaining  $k$  neurons was:

$$f(k; n, p) = \binom{n}{k} p^k q^{(n-k)}, \quad (3)$$

where  $q = 1 - p$ ,  $q$  represents the probability of a retained neuron and  $p$  represents the probability of a neuron turn off,  $p^k q^{(n-k)}$  is the probability of obtaining  $k$  neurons successful sequence in the  $n$  test and  $(n - k)$  failures, while  $\binom{n}{k}$  is the binomial coefficient used to calculate the number of possible successful sequences.

In our lesion segmentation network, we use fixed dropout ratio to handle the feature filtering in each training iteration. The dropout ratio of group TMAX+CBF+CBV is set to 0.01, the dropout ratio of MTT is set to 0.5.

## 2.4 Loss Function

In image segmentation tasks, Dice coefficient (DC) is one of the classic indexes for evaluating the segmentation effect, and it can also be used as a loss function to measure the gap between the result of the segmentation and the ground truth. In binary image segmentation, we use the continuous softmax function outputs to replace the predicted binary labels, we Combine DC with cross entropy function, a pseudo DC loss function proposed in this paper is defined as:

$$L = 1 - \frac{1}{C} \sum_{c=1}^C \left( \frac{2 \sum_{n=1}^N (p(x_n)^c q(x_n)^c)}{\sum_{n=1}^N q(x_n)^c + \sum_{n=1}^N p(x_n)^c} \right), \quad (4)$$

where  $C$  is the class number,  $c \in C$  is the pixel class,  $N$  is the pixel number,  $x_n$  is the  $n$ -th pixel.  $p(x_n)^c$  is a binary value (label) of pixel  $x_n$  belongs class  $c$ , and  $q(x_n)^c$  represents the probability of pixel  $x_n$  predicted by softmax function belongs class  $c$ . In order to measure the loss contribution of each class, aggregating DC from different classes  $C$  as an average. In the traditional single type lesion segmentation task,  $C$  is usually set to 1.

## 3 Experiments and Result

### 3.1 Experiments

We apply MS-DCNN in the ISLES 2018 challenge. The network architecture has shown in Fig. 1, i.e. a dual-pathway DCNN. For ISLES challenge, all CT sequences are resized to  $160 \times 160$ . We use images slices flipped and randomly rotated methods to augment the training images. In training process, the hyper-parameter kept constant: batch size is set to 4, epoch is set to 70, and learning rate is set to 0.001. In our experiment, when the dropout ratio was set as 0.01,

the segmentation results are close to optimal on training dataset. In testing process, network inherits the weight of the training model and realizes the automatic lesions segmentation. After testing, we use the affine transform method to restore the size of all prediction images to the original size. A post-processing step to refine the networks output, we use image median filtering algorithm [14] to alleviate noises and preserve the edge details of images. Finally, we synthesize the 2D slice images into 3D images.

### 3.2 Results

In this challenge, online evaluation is provided with the Dice coefficient (DC) [2], Hausdorff distance (HD) [15], Precision, Recall and AVD as quality metrics. We won't able to see the Ground Truth of the testing dataset. After uploading the segmentation results for the testing dataset, results of each participating team and their ranking be revealed on the challenge websites in a frozen table. We have obtained the scores presented in Table 1.

**Table 1.** The results of our network on ISLES 2018 challenge. Values correspond to the mean (and standard deviation)

Team	DC [0,1]	HD [mm]	Precision [0,1]	Recall [0,1]	AVD [mm]
songt1	0.51 (0.31)	19354856.39 (39507890.41)	0.55 (0.36)	0.55 (0.34)	10.24 (9.94)
pengl1	0.49 (0.31)	19354856.66 (39507890.28)	0.56 (0.37)	0.53 (0.33)	10.08 (10.58)
cheny11	0.47 (0.31)	19384856.55 (39507890.34)	0.56 (0.37)	0.49 (0.33)	11.14 (12.74)
pengl1	0.47 (0.32)	16129055.93 (36779842.04)	0.53 (0.36)	0.47 (0.32)	10.37 (14.42)
Ours	0.44 (0.30)	19354857.71 (39507889.77)	0.54 (0.36)	0.44 (0.32)	10.99 (12.70)

Among the 38 submissions on ISLES 2018, our submission have a superior performance, and ranks fifth. This task is simply too complex and variable for our algorithms to solve. In our training process, our model performs well in segmentation of large lesions. However, smaller and less pronounced lesions are the challenges for our model. As Table 1 shown, compared with DC, Precision and Recall, the values of Hausdorff distance is too high, this may be due to the fact that some lesions are not detected, or there are many outlier points in the our segmentation result. Further work to improve the segment result will consist in optimizing, the particularity of CT image segmentation and incorporating other post-processing to improve the Hausdorff distance.

## 4 Conclusion

In this paper, we proposed the MS-DCNN is an automatic medical image segment network, it surpasses mostly state-of-the-art on ISLES 2018 challenge. Our network inherits previous work and integrates dense blocks. The architecture of U-shape is used to improve the feature locate accurately and semantics

capture. The dense block is used to reuse previous features and alleviate overfitting. In addition, two different dropout rate pathways are used to reduce the number of features between layers and retain important features. Different CT modal sequences play different roles in diagnosis. We will assign different dropout rates to each CT sequence to improve the performance of the current model. At present, our model does not provide precise segmentation for physicians and clinical researchers in this challenge, but it can be used as a support tool.

## References

1. Cardoso, M.J., Sudre, C.H., Modat, M., Ourselin, S.: Template-based multimodal joint generative model of brain data. In: Ourselin, S., Alexander, D.C., Westin, C.-F., Cardoso, M.J. (eds.) IPMI 2015. LNCS, vol. 9123, pp. 17–29. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-19992-4\\_2](https://doi.org/10.1007/978-3-319-19992-4_2)
2. Dice, L.R.: Measures of the amount of ecologic association between species. *Ecology* **26**(3), 297–302 (1945)
3. Erihov, M., Alpert, S., Kisilev, P., Hashoul, S.: A cross saliency approach to asymmetry-based tumor detection. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 636–643. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-24574-4\\_76](https://doi.org/10.1007/978-3-319-24574-4_76)
4. Fakhry, A., Zeng, T., Ji, S.: Residual deconvolutional networks for brain electron microscopy image segmentation. *IEEE Trans. Med. Imaging* **36**(2), 447–456 (2017)
5. Geremia, E., Menze, B.H., Clatz, O., Konukoglu, E., Criminisi, A., Ayache, N.: Spatial decision forests for MS lesion segmentation in multi-channel MR images. *Neuroimage* **57**(2), 378–390 (2011)
6. Glorot, X., Bordes, A., Bengio, Y.: Deep sparse rectifier neural networks. In: International Conference on Artificial Intelligence and Statistics, pp. 315–323 (2011)
7. Gonzalez, R.G., Hirsch, J.A., Koroshetz, W.J., Lev, M.H., Schaefer, P.: Acute ischemic stroke: imaging and intervention. *J. Neuroradiol.* **33**(3), 193 (2006)
8. Grimaud, J., et al.: Quantification of MRI lesion load in multiple sclerosis: a comparison of three computer-assisted techniques. *Magn. Reson. Imaging* **14**(5), 495–505 (1996)
9. Guerrero, R., et al.: White matter hyperintensity and stroke lesion segmentation and differentiation using convolutional neural networks. *NeuroImage: Clin.* **17**(C), 918–934 (2017)
10. Hinton, G.E., Srivastava, N., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.R.: Improving neural networks by preventing co-adaptation of feature detectors. *Comput. Sci.* **3**(4), 212–223 (2012)
11. Hoover, A., Goldbaum, M.: Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels. *IEEE Trans. Med. Imaging* **22**(8), 951–958 (2003)
12. Hoover, A.D., Kouznetsova, V., Goldbaum, M.: Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Trans. Med. Imaging* **19**(3), 203–210 (2000)
13. Huang, G., Liu, Z., Maaten, L.V.D., Weinberger, K.Q.: Densely connected convolutional networks. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 2261–2269 (2017)
14. Huang, T., Yang, G., Tang, G.: A fast two-dimensional median filtering algorithm. *IEEE Trans. Acoust. Speech Signal Process.* **27**(1), 13–18 (1979)



15. Huttenlocher, D.P., Klanderman, G.A., Rucklidge, W.A.: Comparing images using the Hausdorff distance. *IEEE Trans. Pattern Anal. Mach. Intell.* **15**(9), 850–863 (1993)
16. Ioffe, S., Szegedy, C.: Batch normalization: accelerating deep network training by reducing internal covariate shift. [arXiv:1502.03167v3](https://arxiv.org/abs/1502.03167v3)
17. Kamnitsas, K., et al.: Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med. Image Anal.* **36**, 61 (2016)
18. Lcun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. *Proc. IEEE* **86**(11), 2278–2324 (1998)
19. Ledig, C., et al.: Robust whole-brain segmentation: application to traumatic brain injury. *Med. Image Anal.* **21**(1), 40 (2015)
20. Li, X., et al.: 3D multi-scale FCN with random modality voxel dropout learning for intervertebral disc localization and segmentation from multi-modality MR images. *Med. Image Anal.* **45**, 41–54 (2018)
21. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440 (2015)
22. Maier, O., et al.: ISLES 2015 - a public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI. *Med. Image Anal.* **35**, 250–269 (2017)
23. Rao, A., Ledig, C., Newcombe, V., Menon, D., Rueckert, D.: Contusion segmentation from subjects with traumatic brain injury: a random forest framework. In: *IEEE International Symposium on Biomedical Imaging*, pp. 333–336 (2014)
24. Rekić, I., Allassonnire, S., Carpenter, T.K., Wardlaw, J.M.: Medical image analysis methods in MR/CT-imaged acute-subacute ischemic stroke lesion: segmentation, prediction and insights into dynamic evolution simulation models. A critical appraisal. *NeuroImage: Clin.* **1**(1), 164–178 (2012)
25. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) *MICCAI 2015*. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
26. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **15**(1), 1929–1958 (2014)
27. Szegedy, C., et al.: Going deeper with convolutions. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9 (2015)
28. Zhang, R., et al.: Automatic segmentation of acute ischemic stroke from DWI using 3D fully convolutional denseNets. *IEEE Trans. Med. Imaging*, 1 (2018)