Classification of Ischemic Stroke Lesions Based on Cascaded Branch Compression Neural Network

Fengbing JIANG
Science and Technology College of Gannan Normal University, Ganzhou, Jiangxi province, China 341000
Corresponding author’s e-mail: 563622733@qq.com

Abstract. Automatic differentiation of disease images with machine vision technology is of great significance for medical diagnosis. This paper proposes a deep neural network architecture to achieve the classification of ischemic stroke lesions. The proposed architecture uses a cascading approach. The image obtained by down-sampling each image and the original image are respectively input into two convolutional neural networks of different depths and are associated with each other through a cascade structure to solve the defect that the single network cannot balance the local features and the global features. In addition, the use of smaller images obtained after down-sampling can effectively improve the operational efficiency of deep networks. The experimental results show that the proposed algorithm can balance the accuracy and timeliness well.

1. Introduction
Stroke is a serious condition caused by the presence of blood clots or ruptured blood vessels in the blood vessels of the brain. According to statistics from the World Health Organization and other experts, the number of people who lost their lives every year due to cerebral infarction in the world reached 6.2 million [1]. In the past, stroke was considered to be a common disease in the middle-aged and elderly population. But in recent years it has become very common among young people. This has attracted the attention of the whole world. There are two main reasons for stroke. One of them is an ischemic stroke, which is slowly formed due to clogging of certain blood vessels, resulting in insufficient blood supply to the brain. The other is sudden stroke, which affects blood supply to the brain due to rupture of blood vessels in the brain.

This paper proposes a cascaded deep neural network based on brain CT images to distinguish stroke lesions. It is called Cascade Branch Compression Dense Network (CBCDN). The branching architecture is used to fully learn the global and local features in the image and then fuse them to better identify the lesion. In addition, the dual-branch network is separately input images with different pixel sizes, which can shorten the calculation speed of smaller images in a very deep neural network. Larger images are imported into shallower networks to learn more subtle local features.

2. Related Work
The study of stroke is accompanied by a history of modern medical development. Deep learning technology is applied in the field of biomedicine, which is a major contribution to interdisciplinary research. Deep learning technology can achieve strong feature learning ability by using deep neural network. This is different from traditional manual features. In recent years, convolutional neural networks have demonstrated strong capabilities in image classification [2], target detection [3] and
scene analysis [4]. In the field of biomedicine, Maier et al. [5] created a brain stroke recognition image library for research work by relevant personnel, but this is limited to segmentation of lesion areas. This type of public data set is still very scarce due to issues related to ethics and patient privacy. Gao et al. [6] used a convolutional neural network to classify brain MRI images of patients with Olmersheimer's disease, but the network is too complex and does not involve stroke. Padma et al. [7] used the PNN classifier to classify brain CT images, but they focused on the texture features of the brain.

3. Material and Method

DenseNet [8] and other networks work together to implement the proposed cascaded branch compression dense network. That figure 1 shows the complete structure of the Cascade Branch Compression Dense Network (CBCDN).

![Figure 1. Cascade Branch Compression Dense Network (CBCDN)](image)

3.1 Cascade branch compression dense network

DenseNet was proposed by Huang et al. in 2017 and was based on the skip connection structure of ResNet. DenseNet made the skip connection more detailed. A densely connected skip structure is proposed to enable bottom features to be directly input to specific neurons at the top. This minimizes the loss of feature transfer. CBCDN fully absorbs the advantages of DenseNet. It removes the classification layer in the dense structure of the 121 layer and the remaining 120 convolution layers are used as a single branch to train smaller images. The other branch contains only 11 convolutional layers and is paired with 2 pooling layers, which are used to train the original pixel size image.

As is well known, network acceleration technology is an effective strategy to improve the operational efficiency of deep neural networks. The network acceleration methods are mainly divided into three categories, which are downsampled images, downsampling features, and model compression. Downsampling image reduces the size of the input image, which reduces the amount of computation, but loses more features and affects the final accuracy. The downsampling feature is to downsample the feature map output by the middle layer, which still loses more detail features. Model compression adjusts the entire network architecture and is divided into network clipping and pruning. The former directly reduces the number of network layers to reduce parameters. The pruning is to delete the weight links in the network that contribute less to the whole without reducing the number of network...
layers. Direct network clipping can seriously reduce the accuracy, and the degree of discriminating contribution in network pruning is a serious problem. CBCDN combines downsampling images and network clipping techniques to achieve network acceleration more effectively. First, the original image is downsampled to a ratio of 1/2 and then imported into the DenseNet branch. The original pixel size image is input to another branch. 1/2 scale image loses more detail features, but global features are well preserved. DenseNet is able to continuously refine and integrate global features. The original image retains all the most primitive features in the image. He et al. [3] demonstrated that the convolutional layer at the bottom retains most of the fine features, so fewer convolutional layers are used to learn local detail features. However, the pixel size of the DenseNet branch output feature map is 1/16 of the input, i.e., 1/32 of the original image. The other branch outputs 1/16 of the original pixel size. Since feature maps of different pixel sizes cannot be directly fused, a cascade structure is proposed to fuse the bi-branch features and output 1/16 of the original pixel size. The proposed cascade structure is shown in figure 2.

In Fig. 2, F1 represents a feature map of the DenseNet branch output, and F2 represents a feature map of the other branch output. After the F1 input cascade structure, first upsampling to expand its pixel size. Then through a dilated convolution to refine the features. Then perform a batch normalization operation on it and perform a pixel-by-pixel averaging operation with F2. And it is reduced by the convolutional layer and the pooled layer to 1/32 of the original pixel size. Finally, the feature map F3 output by the cascade module is input to the fully connected layer and the Softmax layer for classification.

\[
\begin{bmatrix}
A_{11} & \cdots & A_{im} \\
\vdots & \ddots & \vdots \\
A_{n1} & \cdots & A_{nm_{non}}
\end{bmatrix}
\text{mean}
\begin{bmatrix}
B_{11} & \cdots & B_{1m} \\
\vdots & \ddots & \vdots \\
B_{n1} & \cdots & B_{nm_{non}}
\end{bmatrix}
= \begin{bmatrix}
A_{11}+B_{11} & \cdots & A_{im}+B_{im} \\
\vdots & \ddots & \vdots \\
A_{n1}+B_{n1} & \cdots & A_{nm_{non}}+B_{nm_{non}}
\end{bmatrix}
\] (1)

Where, a feature map of size is represented as a matrix of the same size, and respectively represent an element in the corresponding feature map.

3.2 Dataset and Implementation

As discussed in Chapter 2, CT images of the brain involve issues of ethics and patient privacy. In this study, 319 people from 5 hospitals in China provided CT images of the brain. This includes different individuals suffering from stroke and health. These CT images are visualized into 30,570 gray images. That figure 3 shows some of the images in the collected dataset. It can be found that normal samples and diseased samples are difficult to distinguish unless they are doctors with professional knowledge and experience. For example, the 7th picture in the top line looks very normal, and the 3rd and 5th pictures in the bottom line seem to be sick. Of these images, 21,617 were used for training, 4,395 were used for cross-validation during training, and the remaining 4,558 were used to test the model. All images are unified to a size of 256x256 pixels. Each image is downsampled to the 112x112 pixel size input to the DenseNet branch after inputting the network. The original image is input to another branch.
The model was trained in Tensorflow 1.12 and used 2 NVIDIA GTX 1070 to build a hardware platform. Some of the parameters are set as follows: the batch size is 8, the initial learning rate is 0.001, the learning rate attenuation method uses step, and the optimization method uses the Adam algorithm. Finally, the training and verification results are output after 10 epoch. In addition, accuracy, forward propagation time, and model size are used as indicators to comprehensively evaluate the performance of cascaded branch-compressed dense networks. The training results of advanced architectures such as DenseNet-121, DenseNet-169, ResNet-110 [2], ResNet-164 [9], and WRN [10] are compared with the proposed algorithm.

4. Results and Discussion

The CBCDN trains to obtain the final model after approximately 6 hours, after which the model is tested using the test set. The final results is shown in Table 1. It is worth mentioning that other advanced architectures have different training parameters than the proposed algorithm due to the number of layers or the width of the architecture, such as batch size. But the impact of this change on the results can be ignored, which is the conclusion of many experiments. Only the best results of these advanced architectures are counted in table 1.

As can be seen from the results in Table 1, the proposed algorithm achieves the best accuracy. Compared to the baseline model DenseNet-121, CBCDN has a larger number of parameters because it has 2 branches. However, its forward propagation time is much smaller than the baseline network. This is because its DenseNet branch is being input with a smaller image. In addition, ResNet-110 and ResNet-164 have very few parameters and less computation time, but their accuracy is a big problem. WRN is not optimal in all three indicators. The experimental results fully demonstrate the advancement of the proposed algorithm. It can well realize the classification of stroke lesions in CT images of the brain. That figure 4 shows two convolutional layer weights visualized views of CBCDN model. It can be seen that the weights are very smooth, indicating that the model training is very good. That figure 5 shows the feature maps output by three convolutional layers. The features from the bottom to the top are gradually refined, and the brain features are well captured.
5. Conclusion

From the perspective of human health, this paper proposes a cascaded branch compression dense convolution network to classify stroke lesions in CT images of the brain. This network uses different depths of bi-branched cascaded networks to train sub-graphs of different sizes from the same original image. This played the role of network acceleration. Finally, the classification model is trained using the self-collected brain CT image dataset. It is compared to state-of-the-art architecture to get the highest accuracy. Based on the baseline, the speed of the operation is further improved. The recognition results are still not comparable to experienced doctors. But this technology can be used as a doctor’s aid to reduce the workload of doctors.

References

[1] Mendis S, Puska P and Norrving B 2011 Global atlas on cardiovascular disease prevention and control (Geneva: Geneva World Health Organization)
[2] He K, Zhang X, Ren S and Sun J 2016 Int. Conf. on Computer Vision and Pattern Recognition (Las Vegas) (Washington: IEEE Computer Society) pp 770-778
[3] Ren S, He K, Girshick R and Jian S 2015 Int. Conf. on Neural Information Processing Systems (Montreal) (New York: Curran Associates) pp 91-99
[4] Karlis A D, Kottas T L and Boutalis Y S 2007 J. Elect. Pow. Syst. Res. 77 315-27
[5] Maier O, Menze B H, Gablentz J V D, Levin H, Mattias P H, Matthias L, Stefan W, Abdul B, Paul B, Liang C, Daan C, Francis D, Karl E, Chaolu F, Ben G, Michael G, Tom H, Hanna-Leena H anf Mauricio R 2015 J. Med. Image. Anal. 35 250-69
[6] Gao X W, Hui R and Tian Z 2017 J. Comput. Meth. Prog. Bio. 138 49-56
[7] Padma A, Giridharan N 2016 Int. J. Imag. Syst. Tech. 26 97-105
[8] Huang G, Liu Z, Laurens V D M and Weinberger K Q 2016 Int. Conf. on Computer Vision and Pattern Recognition (Hawaii) (Washington: IEEE Computer Society) pp 2261-69
[9] He K, Zhang X, Ren S and Sun J 2016 Int. Conf. on Computer Vision (Amsterdam) (London: Spinger) pp 630-45
[10] Zagoruyko S, Komodakis N 2016 Wide Residual Networks Preprint gr-qc/0507146