Electronics, Communications and Networks
A.J. Tallón-Ballesteros et al. (Eds.)

doi:10.3233/FAIA231191

Convolutional Neural Network Image Segmentation of Alzheimer's Disease Based on Multi-Order 3D U-NET

Ningsen SUN, Yuan YUAN, Lanhua ZHANG[1], Zhongdong HAN
School of Medical Information and Engineering, Shandong First Medical University and Shandong Academy of Medical Sciences, Tai'an, 271016, China

Abstract. Alzheimer's disease will lead to the atrophy of the Hippocampus. In order to recognize the position changes of Hippocampus, improve the image contour quality, further improve the accuracy of convolution, and achieve the purpose of accurately extracting image information, convolutional neural network is introduced to recognize the Hippocampus region with brain magnetic resonance imaging, and a method combining multi-level 3D U-NET is proposed based on single-stage U-NET. The results showed that this model could enhance the segmentation performance, significantly improved the segmentation accuracy which had certain clinical significance for the brain to recognize the Hippocampus and the automatic discrimination of Alzheimer's disease.

Keywords. Alzheimer's disease, multi-level 3D U-NET, convolutional neural network, image segmentation

1. Introduction

Alzheimer's disease (AD) is a progressive developmental neurodegenerative disease with insidious onset. It is clinically characterized by full-blown dementia manifestations such as memory impairment, visual spatial skill impairment, executive dysfunction, and personality and behavioral changes [1]. As we all known, early detection and early intervention can be based on changes in the hippocampus, and the use of imaging to achieve precise segmentation of AD can more accurately depict the trend of AD [2]. Magnetic resonance imaging (MRI) provides a clearer view of the cerebral spinal cord and nerves, and also distinguishes between grey space and white space [3]. MRI is the imaging modality of choice, and its biomarkers are objectively measurable parameters that offer an important basis as early monitoring and diagnosis of AD. However, due to the complexity of its operation, the excessive level of manpower required for diagnosis, the clinical applications is imminent.

Traditional digital image processing methods are time-consuming and labor-intensive, and are susceptible to various factors, resulting in low accuracy and the possibility of misdiagnosis. Therefore, the application of convolutional neural networks to MRI image segmentation has high clinical value [4]. At present, image segmentation can be divided into region-based methods and contour- and shape-based methods, due

[1] Corresponding Author: Lanhua ZHANG, School of Medical Information and Engineering, Shandong First Medical University and Shandong Academy of Medical Sciences. E-mail: acm_ict@163.com.

to the fact that contour- and shape-based methods need to check various shape information to improve the accuracy of segmentation, so it is relatively difficult to implement. For the region-based method, the boundary energy function is obtained in terms of the statistical grey scale characteristics to generate the relevant contours, which is homogeneous for dealing with separate tissue regions and inhomogeneous for dealing with different tissue regions. For the actual MIR diagnostic process, the difference between the regions is not significant, and the segmentation region is also not obvious, so using the traditional 2D image processing model to deal with the impact of the 3D medicine will be inefficiency and loss of data problems [4]. Given the problems, we invoke the combination of multi-order 3D U-NET model and convolutional neural network, from the k-order 3D U-NET model extends to the k+1-order feature information [5], so as to achieve the depth extraction of image edges based on the convolutional neural network approach to the integration of identification.

In 1998, LECUN [6] applied convolutional neural network to image classification for the first time. In recent years, convolutional neural network had a prominent performance in the field of image segmentation. The image classification through convolutional neural networks which does not need to extract specific features for a specific image dataset can simulate the human brain's visual neural system to automatically screen image features, to achieve personalized classification, and well solve the problem of manually extracting the features of the traditional image classification methods to realize intelligence [7].

In convolutional neural network, the multilayer network structure will be the underlying features of the image through the convolution operation gradually combined to form the high-level features, a multilayer transfer, gradual fusion of the process of feature extraction and classification recognition joint [8].

With the arrival of the artificial intelligence, it is obvious that the application of deep learning in medical diagnosis has the advantage of providing a new direction in order to achieve automatic diagnosis in the clinic, since the proposal of U-NET network. Biomedical images are often composed of many slices to form a whole image. U-NET has become a famous framework for medical image segmentation, and a variety of cases are proving the success of U-NET [9], but a variety of problems are waiting to be solved, such as the 3D image processing problems, over-fitting problems during processing and so on, so it is particularly important to find a suitable processing method. This paper tested a type of convolutional neural network system based on multi-order 3D U-NET, which is used to deal with the early diagnosis of Alzheimer's disease in the clinic.

2. Methods

2.1. Data Sources

In our paper, we used the open source dataset to test our model. The dataset is Kaggle dataset which is from Kaggle group with MRI images. Before the model, all images are resized into 128*128 pixels. The dataset has a total of 6400 images and is divided into four types: Mild Demented, Moderate Demented, Non Demented and Very Mild Demented. Data was collected from several website, hospitals, and public repositories. The sources of the dataset are from webs (https://www.kaggle.com/datasets/jboysen/mri-and-alzheimers; https://catalog.data.gov/dataset/alzheimers-disease-and-healthy-aging-data;) [10].

2.2. Algorithm Model

Convolutional neural network is mainly composed of these types of layers: input layers, convolutional layers, pooling layers, fully connected layers and output layers. By stacking these layers together, we can build a complete convolutional neural network.

Convolutional layers are the central layer for constructing convolutional neural network, and the most important feature is the use of parameter sharing mechanism [11], where the weights of the convolution kernel are obtained through training, and the weights of the convolution kernel will not change in the process of convolution, which indicates that we can extract the same features from different locations of the original image through the operation of a convolution kernel. In brief, although the same target is in different positions in an image, its features are basically the same, so the parameter sharing mechanism effectively reduces the number of training parameters and the risk of over-fitting.

Pooling layers are also known as down sampling layers. Usually a pooling layer is inserted periodically between successive convolutional layers, which serve to gradually reduce the spatial size of the data body. There are many ways of down sampling such as maximum pooling, average pooling and so on. In fact pooling layer can also be seen as a special kind of convolution operation [12].

For any convolution there exists a fully connected layer that implements the forward propagation function. The fully connected layers map the learned feature representations to the labeling space of the samples, integrate them, and act as a classifier in which the final regression analysis model is computed accordingly to get the response value connected to the output layers [13].

After the operation of convolutional and pooling layers, higher level features have been extracted and the output layers take the output of the fully connected layer for further use in classification by normalizing the values and obtaining the probability distribution for each category [14].

In the model, the convolutional layer is (Eq. (1)):

$$\mathrm{K}^2 \times \mathrm{C_i} \times \mathrm{C_o} \times \mathrm{C_o} \tag{1}$$

Where K is the size of the convolution kernel, C_i is the number of input channels, and C_o is the number of output channels. The second term of the arithmetic is the number of parameters of the bias term.

The fully connected layer is (Eq. (2)):

$$T_i \times T_o + T_o \tag{2}$$

Where T_i is the length of the input vector and T_o is the length of the output vector, where the second term is the number of parameters of the bias term.

2.3. Work Flow of Our Works

First of all, we need to do multi-order processing of the data image. In order to achieve the automatic processing of the target contour, we used the 1st order the U-Net network to generate single-order feature information which will be accumulated into the decoding path of the second-order U-Net [15] to generate the second-order feature information, and then repeated the process. There is a corresponding feature layer in the path, and we normalized the image to achieve the multi-order processing of the U-Net. In the process the image is normalized to (0, 1) [16], the MRI image of Alzheimer's disease is used to construct a 3D training, and a convolutional neural network is used to train this network, finally the segmentation process is performed on the target network [17]. The convolutional structure used in the paper is unified as a 3x3 convolutional kernel with 0 padding and 1 striding.

The flowchart is shown in the figure (Fig. 1):

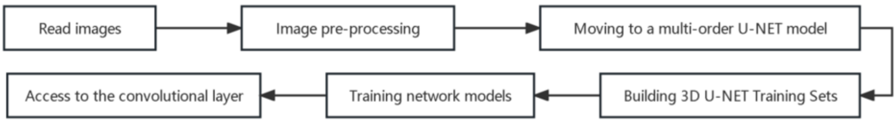

Figure 1. Workflow of the model

3. Results

About the programming, we used Python language to do the processing and test with the version of 3.10 and PyCharm 2022.2.2 in the windows system.

Before analyzing data in the network model, it is necessary to preprocess the data to improve data quality in order to avoid errors in experimental results caused by abnormal data in the dataset. It includes registration, skull removal, image enhancement, tissue segmentation and scale transformation. The image enhancement part includes median filtering and Histogram equalization processing, and the scale transformation includes removing excessive background image volume outside the brain tissue and pruning the image to a pre-selected size for network input 112 × 96 × 96.

The segmentation network consists of an improved U-Net network. The sizes of convolutional kernel, convolution kernel, and maximum pooling are 3 × 3, 2 × 2 and 2 × 2, respectively. ReLU function is used for dividing network, and Sigmoid function is used for activation function.

Before the processing, the images should be preprocessed to the normalization data after some function operations, the preprocessing results are as follows (Fig. 2):

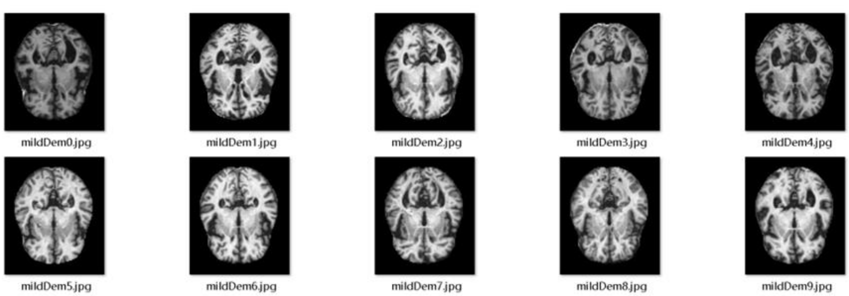

Figure 2. Preprocessing results of imaging

In this model, we adopted the loss function as follows (Eq. (3)):

$$L = 1 - \frac{2\sum_{i}^{N} p_i g_i + smooth}{\sum_{i}^{N} p_i + \sum_{i}^{n} g_i + smooth} \tag{3}$$

L is the loss function, and p is the value of prediction. The super parameter α is 0.7 which is used to balance false positive and false negative rates to alleviate category imbalance issues.

The performance indicator is to verify segmentation performance, and experiments are conducted using the AD dataset. We adopt accuracy, intersection over union (IoU), precision, and recall as evaluation indicators.

This experiment used 179 Alzheimer's MRI images as samples for training. In order to better compare the segmentation effect, data tests were conducted on the brain and background regions of the entire image to obtain the accuracy (mPA = 87.08%), intersection over union (mIoU = 75.94%), precision (mPrecision = 84.38%), and recall rates (mRecall = 87.08%) for different probabilities of brain and background regions.

4. Discussion

This article evaluates the performance indicators of segmentation based on the detection sensitivity, segmentation accuracy, and standard deviation of AD. Assuming that the positive sample is a potential patient of AD and the negative sample is a normal person, it is represented by 1 and 0. For whether the corresponding regional model is accurately divided, the relationship between the predicted value and the true value can be represented by a confusion matrix. Sensitivity is defined as Sensitivity $= \frac{\mathrm{TP}}{\mathrm{TP+FN}}$.

For the image enhance, median filtering belongs to the contrast enhancement part of image enhancement, and its basic principle is to take the median value of the pixel gray level of the field pixel gray level of a pixel point of a gray scale image instead of that pixel point gray level. For images contaminated by pepper noise, median filtering can be used to filter out the noise well while protecting the clarity of the image well. The histogram is the most basic statistical feature of an image, which reflects the distribution of gray values of the image. The purpose of histogram equalization is to homogenize the distribution of the image over the entire dynamic range of gray values, improve the brightness distribution state of the image, and enhance the visual effect of the image.

For the test set, the image is divided into 2 × 2 small blocks in the large pooling layer. In each block, the maximum value is returned to the output, and the 4 × 4 image is pooled to become a 2 × 2 image.

An image tensor of (1, 28, 28) is input into a convolutional network, passing through a convolutional layer with a kernel size of (5, 5) and an output channel of 10. After passing through a maximum pooling layer of (2.2) and a convolutional layer with a kernel size of (5.5) and an output channel of 20, the channel changes from 10 to 20. Then, it passes through a maximum pooling layer of (2.2) and finally outputs as a fully connected layer of 10 dimensions through a 320, Obtain ten probability values corresponding to ten types.

In order to improve the stability of the scheme, three independent running experiments were conducted in this experiment, with accuracy rates of 98.31%, 98.72%, and 98.32%, respectively, with an average accuracy rate of 98.45%. This result effectively indicates that the method proposed in this article is effective. Meanwhile, for accuracy, intersection over union, accuracy, and recall rate, the results of brain and background regions were different, their corresponding values are 0.81 to 0.93, 0.62 to

0.90, 0.73 to0.96, 0.81 to 0.93, which implied that the background regions were better than brain regions.

The model in this paper is compared with the models in the remaining four papers. Specifically a dual-attention 3D-UNet based segmentation network model proposed by Wang et al. [18] An improved U-Net segmentation method proposed by Xiao et al. [19] A segmentation network model for segmentation adversarial networks proposed by EI-regaily et al. [20] A framework for convolutional neural networks proposed by Mukherjee et al. [21] The specific experimental data is shown in the Tab. 1.

Table 1. Scores in models

Model	mPA	mIOU	mPrecision	mRecall
Wang	——	79.40	——	82.10
Xiao	83.74	——	——	84.10
EI-regaily	80.14	——	66.10	——
Mukherjee	59.40	52.70	——	——
Our method	87.08	75.94	84.38	87.08

The experimental results show that the network model proposed in this paper is better than the above methods in general, proving the feasibility and effectiveness of the model.

5. Conclusions

The project integrates the improvement of 3tunet network into the convolutional neural network, uses the thinking of multi-level eigenvalue to highlight the image area and make it more accurate for the next step of convolutional neural network. The combination of the two enables it to accurately segment the brain image of AD, solves the problem of gradient explosion, and improves the accuracy of testing.

At present, there are still some weaknesses, such as excessive memory usage and cumbersome processing. Early diagnosis of AD can have a certain early intervention effect on later diseases, which has important clinical added value for brain image segmentation processing. Therefore, in order to conduct in-depth research, reduce computational complexity, use more clinical data, and increase the possibility of clinical application.

Acknowledgements

This research was supported by the Students' project for innovation and entrepreneurship training program of Shandong First Medical University and Shandong Academy of Medical Sciences (No. 2022104391008), Youth Science Foundation Cultivation Support Program of Shandong First Medical University and Shandong Academy of Medical Sciences (No. 202201-036) and the Tai'an Science and Technology Innovation Development Project (No. 2022GX025).

References

[1] Matharu B, Spencer N, Howe F, et al. Gadolinium-complexed Aβ-binding contrast agents for MRI diagnosis of Alzheimer's Disease. Neuropeptides, 2015. DOI:10.1016/j.npep.2015.07.001.

[2] Wells AJ, Holmes EH, O'Callaghan MJ, et al. Increased Cerebral Vascular Reactivity in the Tau Expressing rTg4510 Mouse: Evidence against the Role of Tau Pathology to Impair Vascular Health in Alzheimer's Disease. Journal of Cerebral Blood Flow & Metabolism, 2015, 35(3): 359-62.

[3] Jack CRJ, Petersen RC, Xu YC, et al. Prediction of AD with MRI-based hippocampal volume in mild cognitive impairment. Neurology, 1999, 52(7):1397-1403.

[4] Rayhan MA, Ferdous AA, Uddin RA, et al. DoubleU-NetPlus: a novel attention and context-guided dual U-Net with multi-scale residual feature fusion network for semantic segmentation of medical images. Neural Computing and Applications, 2023, 35(19): 14379–14401.

[5] Wang L, Zhang G. Cluster Ensemble Based Image Segmentation Algorithm. International Journal of Advanced Robotic Systems, 2013, 10(4):1.DOI:10.5772/56769.

[6] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 1998, 86(11):2278-2324.

[7] Do-Hai-Ninh N, Minh-Nhat T, Viet-Dung N, et al. An EffcientNet-encoder U-Net Joint Residual Refinement Module with Tversky–Kahneman Baroni–Urbani–Buser loss for biomedical image Segmentation. Biomedical Signal Processing and Control, 2023,83: 104631. https://doi.org/10.1016/j.bspc.2023.104631.

[8] Kishore KR, Rao KS. A novel approach to unsupervised pattern discovery in speech using Convolutional Neural Network.Computer Speech & Language, 2022, 71(1):101259.

[9] Zhang Y, Zhong P, Jie D, et al. Brain Tumor Segmentation From Multi-Modal MR Images via Ensembling UNets, Frontiers in Radiology, 2021(1):704888.

[10] Sachin Kumar, Sourabh Shastri. Alzheimer MRI Preprocessed Dataset. Kaggle. 2022. https://www.kaggle.com/datasets/sachinkumar413/alzheimer-mri-dataset.

[11] Teresa MG, Héctor A, Alberto JB, et al. Sentiment analysis in non-fixed length audios using a Fully Convolutional Neural Network. Biomedical Signal Processing and Control, 2021,69(2):102946.

[12] Rajee MV, Mythili C. Gender classification on digital dental x-ray images using deep convolutional neural network. Biomedical Signal Processing and Control, 2021, 69(1):102939.

[13] Niyas S, Vaisali SC, Show I, et al. Segmentation of focal cortical dysplasia lesions from magnetic resonance images using 3D convolutional neural networks. Biomedical Signal Processing and Control, 2021, 70:102951.

[14] Yang D, Lan L, Lei Y, et al. Automated CT pancreas segmentation for acute pancreatitis patients by combining a novel object detection approach and U-Net. Biomedical Signal Processing and Control, 2023,81: 104430.

[15] Xue Feng, Nicholas Tustison, Craig Meyer. Brain Tumor Segmentation using an Ensemble of 3D U-Nets and Overall Survival Prediction using Radiomic Features. Computer Vision and Pattern Recognition, 2018. arXiv:1812.01049.

[16] Yimin C, Yuqing L, Zhenggong H, et al. Swin Unet3D: a three-dimensional medical image segmentation network combining vision transformer and convolution. BMC medical informatics and decision making, 2023, 23(1). DOI: 10.1186/s12911-023-02129-z.

[17] Sarah Montagne, Dimitri Hamzaoui, Alexandre Allera, et al. Challenge of prostate MRI segmentation on T2-weighted images: inter-observer variability and impact of prostate morphology. Insights Imaging, 2021, 12(1):71. DOI: 10.1186/s13244-021-01010-9.

[18] Wang P, Qiang Y, Yang XT et al. Lung nodule segmentation network model based on dual-attention 3D-UNet. Computer Engineering, 2021, 47(02):307-313.

[19] Xiao N, Qiang Y, Zhao J et al. Lung nodule segmentation based on segmentation adversarial network. Computer Engineering and Design, 2019, 40(04):931-936+1076.

[20] EL-REGAILY SA, SALEM MAM, AZIZ MHA, et al. Lung nodule segmentation and detection in computed tomography. 2017 Eighth International Conference on Intelligent Computing and Information Systems, 2017:72-78.

[21] MUKHERJEE S, HUANG X, BHAGALIA RR. Lung nodule segmentation using deep learned prior based graph cut. 2017 IEEE 14th International Symposium on Biomedical Imaging, 2017:1205-1208.