Transactions on Computational and Scientific Methods | Vo. 4, No. 12, 2024
ISSN: 2998-8780
https://pspress.org/index.php/tcsm
Pinnacle Science Press

Fully Convolutional Neural Networks for High-Precision Medical Image Analysis

Zitao Zheng[1], Yanlin Xiang[2], Yijiashun Qi[3], Yang Lin[4], Hanchao Zhang[5]
[1]Independent Researcher, New Jersey, USA
[2]University of Houston, Houston, USA
[3]University of Michigan, Ann arbor, USA
[4]University of Pennsylvania, Philadelphia, USA
[5]New York University , New York, USA
*Correspondence should be addressed to Yang Lin**;** yang.lin1345@gmail.com

Abstract:

This study proposed a medical image semantic segmentation algorithm based on a fully convolutional neural network (FCNN) for accurate segmentation of anatomical structures and lesion areas in medical images. In the experiment, we used mIOU and Dice coefficients as the main evaluation indicators to verify the superior performance of FCNN in fine-grained segmentation tasks. The experimental results show that FCNN performs well in both mIOU and Dice values, surpassing traditional semantic segmentation models such as U-Net, SegNet, and DeepLabV3. This shows that the convolutional structure and end-to-end feature extraction mechanism of FCNN can effectively capture the complex boundaries and subtle features in medical images and achieve high-precision pixel-level segmentation. The algorithm has good generalization and strong cross-modal adaptability, and is suitable for a variety of medical image types such as CT and MRI. In the future, with the development of deep learning technology, the application potential of FCNN in the field of medical image analysis will be further expanded, providing technical support for precision medicine and intelligent diagnosis and treatment, and providing auxiliary decision-making bases for medical workers.

Keywords:

FCNN, medical image segmentation, semantic segmentation, deep learning

1. Introduction

In medical image analysis, semantic segmentation is particularly critical. It can help medical personnel identify specific anatomical structures or lesion areas by accurately dividing different areas in the image [1]. Fully Convolutional Neural Networks (FCNN), as an advanced deep learning technology, has shown great potential in the semantic segmentation of medical images. FCNN replaces the fully connected layers in the traditional convolutional network with convolutional layers and deconvolutional layers, allowing the network to process input images of any size and generate segmentation maps of corresponding sizes. This design not only improves the accuracy of segmentation but also reduces the number of network parameters, thereby improving computational efficiency. The end-to-end learning capability of FCNN enables it to automatically complete feature extraction and segmentation tasks from input to output, providing strong technical support for the accurate segmentation of medical images [2].

Another advantage of FCNN is its highly automated end-to-end segmentation capability. Traditional image segmentation methods usually require multiple steps, including image preprocessing, feature extraction, region segmentation, and post-processing, while FCNN can integrate these steps into a unified framework, thereby reducing manual intervention [3]. This end-to-end architecture not only improves the efficiency of segmentation, but also avoids the error accumulation caused by multiple steps, which helps to improve the overall accuracy of segmentation[4]. Especially when facing complex anatomical structures or irregular lesions, FCNN can capture subtle morphological features through the multi-layer feature extraction capability of deep learning, thereby achieving a more refined segmentation effect[5]. This highly automated segmentation capability has made FCNN widely recognized in the application of medical image processing, especially for the efficient processing of large-scale medical image data.

In addition, FCNN has strong cross-modal adaptability, which makes it superior in different types of medical images[6]. Whether it is CT, MRI or ultrasound imaging, the convolutional structure of FCNN can adapt to the feature extraction requirements under different imaging modalities and achieve stable segmentation effects[7]. Since medical data usually contains multiple imaging modalities, and various modalities have large differences in image features, resolution and noise, FCNN has shown high generalization in different imaging modalities with its diverse understanding and processing capabilities of feature levels. This cross-modal adaptability not only improves the practicality of the model, enabling it to cope with the diverse needs in different clinical environments, but also lays the foundation for the cross-modal application of medical image analysis[8].

FCNN also has outstanding performance in improving segmentation accuracy. Its pixel-level prediction ability enables the model to accurately distinguish the boundary between the lesion area and normal tissue, thereby improving the meticulousness of segmentation. This is of great significance for accurate diagnosis, treatment planning and surgical navigation. In the segmentation process of FCNN, the multi-layer convolution structure extracts image features layer by layer, which can gradually enhance the model's ability to distinguish subtle structures, thereby achieving higher accuracy in lesion segmentation. Especially in small lesions or lesion areas with blurred boundaries, FCNN can more accurately capture the morphology and location of lesions through comprehensive analysis of deep features, providing doctors with clearer and more reliable segmentation results, and laying a solid foundation for the formulation of personalized treatment plans.

In general, the semantic segmentation algorithm of medical images based on FCNN has significant advantages in automated processing, cross-modal adaptability and high-precision segmentation. With the continuous advancement of deep learning and computer hardware technology, the application prospects of FCNN in medical image processing are broader, especially in the fields of disease screening, early diagnosis and treatment monitoring. In the future, FCNN is expected to be widely used in more detailed medical scenarios, providing important technical support for the development of smart medicine, and contributing to improving the quality of medical services, improving diagnosis and treatment efficiency, and promoting precision medicine.

2.Method

In this study, we use a fully convolutional neural network (FCNN) to achieve the semantic segmentation task of medical images. The core of FCNN lies in the fully convolutional structure, which allows the model to directly generate pixel-level segmentation results from the input image. Its network architecture is shown in Figure 1.

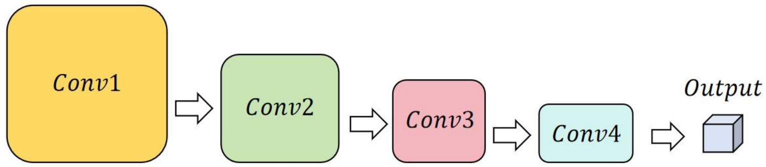

Figure 1. Overall architecture of the model

Traditional convolutional neural networks usually add a fully connected layer to the last layer, while FCNN replaces the fully connected layer with a convolutional layer to preserve spatial information. Specifically, given an input image X, a feature map F is obtained after multiple layers of convolution operations. In each layer of convolution operation, the feature extraction process can be expressed as:

$$F_l = f(W_l * F_{l-1} + b_l)$$

In order to achieve pixel-level segmentation, FCNN uses deconvolution (or upsampling) operation after feature extraction to restore the low-resolution feature map to the original resolution of the input image. The upsampling process can be achieved through deconvolution operation, that is, converting feature map A into segmentation map B, where the position of each pixel represents the segmentation category. The formula for the deconvolution operation is:

$$S = W_{deconv} * F + b$$

Among them, W_{deconv} is the deconvolution kernel and b is the bias term. Through this operation, the model can predict each pixel, thereby outputting the category label to which each pixel belongs and completing accurate segmentation.

In order to optimize the model, we use the cross-entropy loss function during training to measure the difference between the predicted result and the true label. Assuming that the prediction output of the model is y_i and the true label is y'_i, the loss function is defined as:

$$L = -\frac{1}{N}\sum_{i=1}^{N}(y_i \log(y'_i) + (1-y_i)\log(1-y'_i))$$

Among them, N represents the total number of samples, y_i and y'_i are the true label and predicted probability of the i-th sample respectively. By minimizing the loss function, the model can continuously adjust the parameters to make the segmentation results more accurate.

In practical applications, in order to improve the stability and convergence speed of FCNN, we also used batch normalization and Dropout techniques, which effectively reduced overfitting and improved the generalization ability of the model. Finally, FCNN can output accurate segmentation maps, providing fine segmentation results for medical image analysis.

3. Experiment

3.1 Datasets

In this study, we used a real dataset widely used in medical image analysis, the "2017 AAPM Thoracic CT Segmentation Challenge Dataset", which is a chest CT image dataset provided by the American Association of Physics in Medicine (AAPM). This dataset contains hundreds of chest CT scans from different patients, mainly used for automatic segmentation of lung structures. The dataset provides detailed labels, including segmentation masks of different anatomical structures such as lungs, airways, and lobes, making it ideal for studying semantic segmentation tasks.

Each CT image is normalized and contains expert annotations of lung regions and other important anatomical structures. The images in the dataset have high resolution and contain rich details, which can be effectively used to train deep learning models, especially for fine-grained segmentation tasks. The dataset also contains multi-angle image slices of different patients, allowing the model to learn and extract features of lung structures from multiple perspectives. In addition, the dataset also provides detailed patient metadata, such as age, gender, and health status, to facilitate further hierarchical analysis and personalized model training.

In the experiment, the dataset was divided into training set, validation set, and test set to ensure the generalization ability of the model on different data. Through this dataset, the FCNN model can learn complex anatomical structures in real scenes, identify the boundaries and shape features of the lungs and airways, and provide high-quality data support for medical image semantic segmentation tasks. Due to the richness of the data and the accuracy of the annotations, this dataset is not only suitable for the diagnosis and treatment of lung diseases, but also provides a reliable foundation for the development of intelligent medical systems.

3.2 Experimental setup

In the experimental setting, first, we divided the dataset into training set, validation set and test set, with the proportions of 70%, 15% and 15% respectively. In order to ensure the generalization ability of the segmentation model on different patients and different image slices, we ensured that the images in each subset had similar distribution characteristics when dividing the data. The training set is used to learn the model parameters, the validation set is used to adjust the model hyperparameters to prevent overfitting, and the test set is used for the final performance evaluation to ensure that the model can obtain reliable segmentation effects on unseen data.

During the model training process, we used the fully convolutional neural network (FCNN) architecture to achieve pixel-level semantic segmentation. To optimize the model parameters, we used the Adam optimizer with the initial learning rate set to 0.001 and a dynamic learning rate decay strategy to improve the convergence speed and stability. The batch size was set to 8 to adapt to the video memory requirements and ensure computational efficiency. During training, we also applied data augmentation techniques, including random rotation, translation, scaling and flipping, to improve the robustness and generalization ability of the model, so that it can adapt to medical images of different angles and scales.

In order to evaluate the segmentation performance of the model on the test set, we use the Dice coefficient and IoU (intersection over union) as the main evaluation indicators. The Dice coefficient can measure the overlap between the model segmentation results and the true label, while IoU evaluates the similarity between the model prediction area and the actual area. These two indicators can fully reflect the performance of the model in terms of boundary accuracy and segmentation completeness. Through these experimental settings, we ensure that the FCNN model can obtain stable and accurate segmentation results in the task of semantic segmentation of medical images, providing strong support for practical medical applications.

3.3 Experiments

In this experiment, we compared five commonly used semantic segmentation models to evaluate the performance of FCNN in medical image segmentation. These models include U-Net, SegNet, DeepLabV3, PSPNet, and Mask R-CNN. U-Net is an encoder-decoder structure designed for biomedical images, which effectively retains feature information through skip connections; SegNet is based on the VGG architecture and uses layer-by-layer upsampling, which is suitable for multi-scale image segmentation; DeepLabV3 uses dilated convolution to expand the receptive field, thereby improving the ability to capture details; PSPNet obtains contextual information of different scales through the spatial pyramid pooling module, which is suitable for segmentation tasks with complex backgrounds; Mask R-CNN adds pixel-level mask prediction on the basis of Faster R-CNN, which can perform object detection and segmentation at the same time. Through the comparative experiments of these models, we can comprehensively analyze the segmentation advantages of FCNN under different feature extraction mechanisms, and further verify its applicability in the task of semantic segmentation of medical images. The experimental results are shown in Table 1.

Table 1. Experiment result

Model	mIOU%	Dice
U-Net	78.3	80.1
SegNet	76.5	78.9
DeepLabV3	80.7	82.3
PSPNet	81.2	83.0
Mask R-CNN	82.1	83.7
FCNN(Ours)	84.5	86.2

From the experimental results in Table 1, we can see that FCNN performs best in the task of semantic segmentation of medical images, surpassing other models with 84.5% mIOU and 86.2% Dice coefficient. This result shows that FCNN has stronger capabilities in fine-grained image segmentation. mIOU (mean intersection over union) and Dice coefficient are core indicators for measuring segmentation accuracy, reflecting the degree of overlap between the predicted area and the actual area and the overall accuracy of the model, respectively. FCNN can achieve the best results due to its efficient convolutional structure and targeted optimization, which enables the model to capture more detailed boundary information while maintaining a high resolution, especially suitable for complex medical image segmentation tasks.

As a traditional medical image segmentation network, U-Net achieved 78.3% mIOU and 80.1% Dice coefficient, which is slightly lower than deeper models, but still performs well. The advantage of U-Net lies in its encoding-decoding structure and skip connections, which allow features to be transferred between different scales, so it has good generalization ability in small sample and fine-grained medical image tasks. Despite this, U-Net has certain limitations in capturing complex boundaries and deeper features, so it failed to achieve higher scores in mIOU and Dice, but it still performed stably when dealing with relatively simple segmentation tasks such as anatomical structures.

DeepLabV3 and PSPNet also showed good segmentation capabilities in the experiment, reaching 80.7% and 81.2% mIOU, and 82.3% and 83.0% Dice coefficients respectively. DeepLabV3 introduces dilated convolution to expand the receptive field, enhances the extraction of local features while maintaining resolution, and can effectively identify complex boundary areas, so it is better than U-Net and SegNet in segmentation accuracy. PSPNet further strengthens the fusion of contextual information at different scales through the spatial pyramid pooling module. This multi-scale feature fusion method enables PSPNet to have better segmentation effects when dealing with images with complex backgrounds and blurred boundaries. However, due to their complex structures, these two are slightly inferior to FCNN in processing boundaries and details of medical images.

Mask R-CNN also performed well in mIOU and Dice coefficient, reaching 82.1% and 83.7% respectively. Mask R-CNN was originally used for target detection tasks. By adding pixel-level segmentation capabilities on the basis of Faster R-CNN, it has strong multi-task processing capabilities. Its advantage is that it can simultaneously identify objects and perform fine segmentation, so it performs well in medical images with multiple targets and overlapping areas. However, since Mask R-CNN mainly focuses on the overall detection of objects, the accuracy of segmentation boundaries is slightly lower than that of FCNN optimized for semantic segmentation. This leads to its slight shortcomings in the processing of details in medical images, especially for complex anatomical structures.

In summary, FCNN performed best in this experiment, indicating that its design is particularly suitable for fine-grained semantic segmentation of medical images. Its efficient end-to-end architecture and optimized convolutional feature extraction method enable it to generate high-precision segmentation results in a short time. This is crucial for applications in the medical field. The outstanding performance of FCNN provides accurate data support for disease diagnosis, treatment plan formulation, and surgical navigation. In the future, with the improvement of deep learning and computing power, FCNN is expected to be further optimized to provide stronger technical support for more complex medical image segmentation tasks.

Finally, we present the graphs of the evaluation indicators increasing with epoch, and the results are shown in Figures 2 and 3.

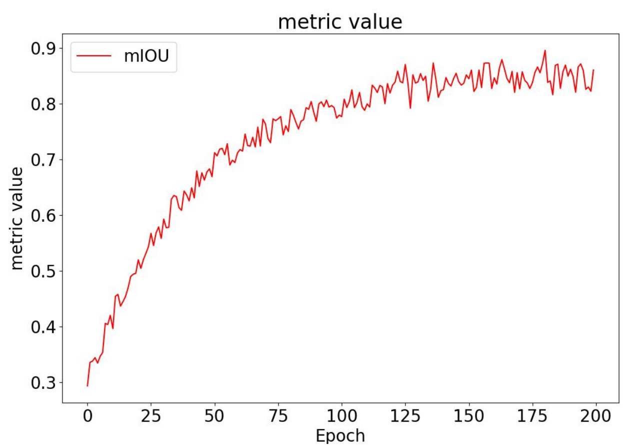

Figure 2. Image of mIOU increasing with epoch

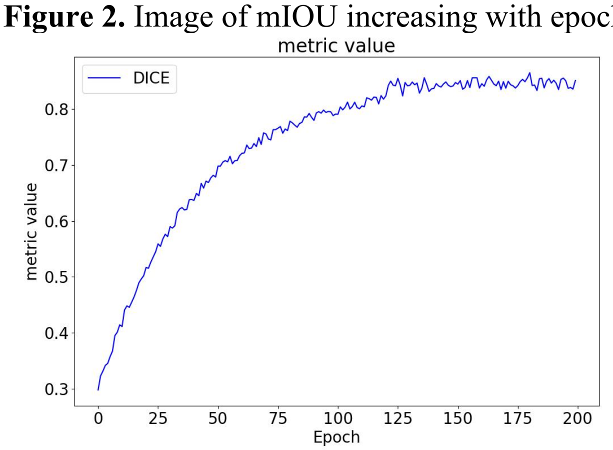

Figure 3. Image of DICE increasing with epoch

Figures 2 and 3 show the changes in two core indicators of the model during training: mIOU (mean intersection over union) and Dice coefficient. The horizontal axis represents the training cycle (Epoch), and the vertical axis represents mIOU and Dice value respectively. It can be observed from the figure that as the training progresses, both indicators show a continuous upward trend, indicating that the model is gradually learning and improving its performance in image segmentation tasks.

In the first figure, mIOU gradually increases from the initial value, and the curve gradually stabilizes after about 100 epochs. This shows that the model quickly learns and optimizes the main features in the early stage. After reaching a certain level of training, the improvement rate of mIOU slows down significantly, and finally stabilizes at about 0.85 when it is close to 200 epochs. This trend shows that the model has fully learned the feature information in the data, and the segmentation accuracy has reached a high level.

The second figure shows the change of the Dice coefficient. The Dice value also gradually increases from the initial value, and stabilizes in the later stage of training, and finally stabilizes at about 0.88. The rising trend of the Dice coefficient is consistent with that of mIOU, indicating that the model performs well in terms of accuracy and overlap. The small fluctuation of the Dice coefficient indicates that the model is relatively stable in boundary and detail segmentation. After multiple iterations, the segmentation results of the model are close to the optimal state.

Overall, these two indicators continue to rise during the training process and tend to stabilize in the later stage, showing that the model has good convergence and has achieved the best segmentation effect. This stable improvement trend shows the learning effect of the model and the effectiveness of the optimization process, providing strong technical support for high-precision segmentation in practical applications.

4.Conclusion

Through the experimental results of this study, we verified the superior performance of the FCNN-based medical image semantic segmentation algorithm in fine-grained segmentation tasks. Through the improvement of mIOU and Dice coefficients, we found that FCNN can effectively capture the complex boundaries and subtle features in medical images and achieve high-precision pixel-level segmentation. This result shows that the convolution structure and feature extraction mechanism of FCNN are particularly suitable for the semantic segmentation of medical images, providing accurate data support for clinical applications such as disease diagnosis and surgical planning.

Although the experimental results of this study have shown a high segmentation effect, the generalization ability and real-time performance of the model are still important research directions in actual medical applications. FCNN has a large amount of calculation in the training and inference process, and there is still room for optimization for application scenarios with high real-time processing requirements. In the future, the real-time performance of FCNN can be improved by introducing a lightweight model structure, reducing the number of parameters and improving hardware performance, so that it can adapt to more medical scenarios with high-efficiency requirements.

Looking forward to the future, with the further development of deep learning and computer hardware technology, FCNN and its improved models will play an important role in more medical image processing tasks. We can expect that combined with new technologies such as cross-modal data fusion and reinforcement learning, FCNN will show broader application potential in disease prediction, personalized treatment, and health management, providing solid technical support for the realization of intelligent and precise medical services.

References

[1] Singh, Pranav, and Jacopo Cirrone. "Exploring Intrinsic Properties of Medical Images for Self-Supervised Binary Semantic Segmentation." arXiv preprint arXiv:2402.02367 (2024).

[2] Hu, Hengqi, et al. "Deep Dual-Stream Convolutional Neural Networks for Cardiac Image Semantic Segmentation." IEEE Transactions on Industrial Informatics (2024).

[3] Awasthi, Aishwary, Ramesh Chandra Tripathi, and T. Thiruvenkadam. "Real-Time Semantic Segmentation of Medical Images Using Convolutional Neural Networks." 2024 International Conference on Optimization Computing and Wireless Communication (ICOCWC). IEEE, 2024.

[4] Al-Qurri, Ahmed, and Mohamed Almekkawy. "Ultrasound Image Segmentation using a Model of Transformer and DFT." 2024 IEEE UFFC Latin America Ultrasonics Symposium (LAUS). IEEE, 2024.

[5] Chen, Jialu, and Baohua Yuan. "CCFNet: Collaborative Cross-Fusion Network for Medical Image Segmentation." Algorithms 17.4 (2024): 168.

[6] Feng, Yanbo, Adel Hafiane, and Hélène Laurent. "A weakly supervised end-to-end framework for semantic segmentation of cancerous area in whole slide image." Pattern Analysis and Applications 27.2 (2024): 35.

[7] Wei, Zhenkun, Jia Liu, and Yu Yao. "Semantic-Based Optimization of Deep Learning for Efficient Real-Time Medical Image Segmentation." International Journal on Semantic Web and Information Systems (IJSWIS) 20.1 (2024): 1-16.

[8] Rayachoti, Eswaraiah, Ramachandran Vedantham, and Sanjay Gandhi Gundabatini. "EU-net: An automated CNN based ebola U-net model for efficient medical image segmentation." Multimedia Tools and Applications (2024): 1-25.