

Review of Medical Image Segmentation Methods

Mengjie Gu, Qingtao Wu

Henan University of Science and Technology, Luoyang 471000, China

Abstract: Medical image is one of the key factors in the process of medical diagnosis and treatment. By analyzing the medical image obtained, doctors make judgments on the patient's condition and plan the next treatment process. Medical image segmentation is the process of segmenting areas of interest in medical images according to specific needs, which is a key step in medical image processing and analysis. With the great improvement of computer processing power, how to quickly and effectively segment huge image data and mining valuable information is the research hotspot of segmentation algorithm at present. This paper introduces several traditional medical image segmentation methods such as threshold method and region method, as well as convolutional neural network and Transformer segmentation method based on deep learning technology. Finally, this paper makes a summary and prospects for future development.

Keywords: Medical Image Segmentation; Traditional Methods; Deep Learning.

1. Introduction

With the continuous development of modern medical imaging equipment, the diagnosis of clinical diseases is closely related to the processing of medical images. As shown in Figure 1, focal areas can be clearly identified through imaging technology. However, in clinical medicine, doctors mainly rely on subjective judgment to diagnose diseases. In a high-intensity working atmosphere, it is very likely to occur that different doctors at the same time have inconsistent diagnosis of the same image or the same doctor has inconsistent diagnosis of the same image at different times.

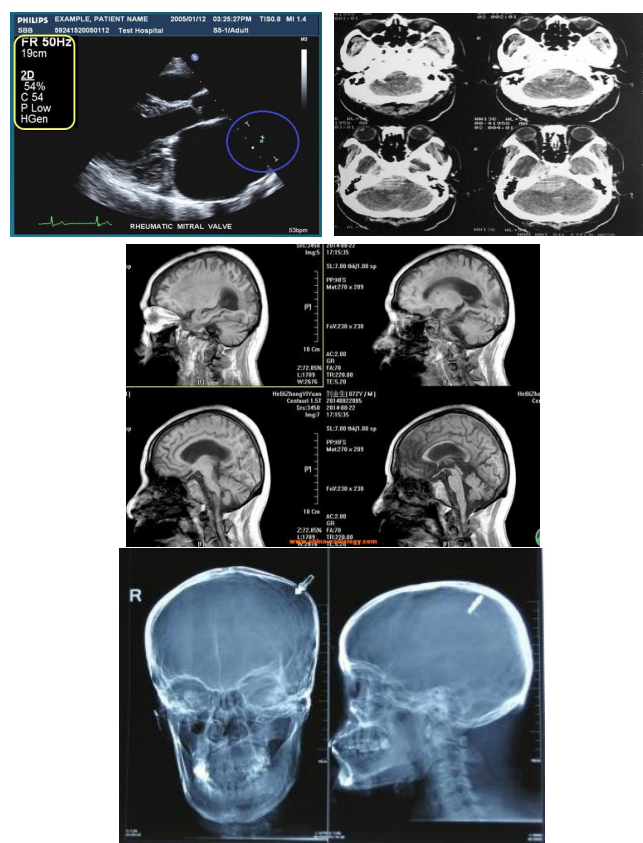


Figure 1. Imaging with multiple medical devices.

In order to improve the diagnostic efficiency and reduce the pressure on doctors, computer-assisted diagnosis and treatment [1] has become a research hotspot in the field of intelligent medicine at the present stage. The combination of computer and medical field has created a new era of digital medical diagnosis. Medical image processing mainly includes image detection [2-3], image segmentation [4-5], image registration [6-7], image fusion and other aspects. Among them, medical image segmentation plays an important role in medical image processing, through this technology can extract lesion tissue and measure its size to observe the disease, directly related to the accuracy of computer medical assisted diagnosis system. At the same time, image-guided surgery, tumor radiotherapy and other technologies are also developed on the basis of accurate segmentation of medical images, which provides a quantitative and qualitative basis for doctors to diagnose diseases.

However, medical image segmentation is a complicated and difficult task, so far no perfect solution has been obtained. Compared with ordinary images, medical images are characterized by fuzziness, complexity and individual difference. In addition, segmentation algorithms are usually oriented to specific segmentation tasks, and there is no general algorithm that can perfectly perform effective segmentation for all medical images. At present, medical image segmentation algorithms based on traditional methods and medical image segmentation based on deep learning are more commonly used. This paper summarizes the two medical image segmentation methods.

2. Traditional Medical Image Segmentation Method

2.1. Threshold Techniques

In the medical image segmentation method based on threshold [8], a global threshold T is set firstly for image $f(x, y)$. If $f(x, y) > T$, it is called the foreground point (x, y) ; otherwise, it is called the background point. The formula is as follows:

$$g(x, y) = \begin{cases} 1, & f(x, y) > T, \\ 0, & f(x, y) \leq T, \end{cases} \quad (1)$$

If there are multiple objects in the image, it can be

segmented by setting multiple global thresholds. When the background lighting is not uniform or there is noise in the image, if the global threshold processing is usually unable to achieve effective segmentation, then the variable local threshold can be used to divide the image into several image blocks, and set a threshold for each image block.

The segmentation method based on threshold has the advantages of small computation, simple implementation and high efficiency, which is one of the most common means of medical image segmentation direction. However, it also has some limitations. For example, this method is only applicable to some images with simple targets and large differences in background gray values, and is sensitive to noise, so its robustness is not high.

2.2. Region Segmentation

The region segmentation method [9] is a segmentation technology for the purpose of finding regions, including two basic forms: one is the region growth [10]; The other is region split and merge [11].

The region growth is the process of combining pixels or subregions into larger areas according to pre-defined growth criteria. This method needs to solve three problems: 1) Select seed pixels that can accurately represent the desired region; 2) Determine the criteria for adding adjacent pixels to seeds during growth; 3) Set the conditions for stopping the growth process. The region growth method is simple to calculate and has a good segmentation effect on connected regions with the same characteristics. However, for images with noise, over segmentation is highly likely to occur.

Region splitting and aggregation can be said to be the inverse process of region growth, which subdivides an image into a set of arbitrarily disjoint regions and then aggregates or splits these regions. The details are as follows: let R represent the whole image, and select an attribute Q , if $Q(R_i) = FALSE$, then the image will be divided into four disjoint quadrant regions; If no further division is possible, any two adjacency areas R_j and R_k that meet $Q(R_j \cup R_k) = TRUE$ are aggregated. Stop the operation when no further aggregation is possible. Compared with threshold method, region splitting and aggregation method have better segmentation effect on complex image, but their algorithms are more complex and require a large amount of calculation, so the region boundary may be destroyed when splitting.

In medical image applications, region growing and region splitting and merging algorithms are usually combined. Such algorithms have a better effect on the segmentation of complex scenes defined by some complex objects or the segmentation of some natural scenes with insufficient prior knowledge.

2.3. Graph Theory Segmentation

Medical image segmentation based on graph theory [12-13] abstracts the image to be divided into a weighted graph $G = (V, E)$, pixel points in the image are node V of the weighted graph, and the relationship between pixel points is edge E of the weighted graph, whose weight $W(i, j)$ represents the similarity between vertex i and vertex j in space, texture, color and other aspects. The weighted graph is cut into subsets according to the degree of similarity. The pixel similarity within the subset is high, while the pixel similarity between subsets is low. The cutting process is actually to remove the edges between nodes with low similarity, so as to achieve segmentation.

Graph theory segmentation is flexible and requires less prior knowledge, but it cannot segment very similar objects with unclear boundaries, which may result in wrong segmentation. In medical applications, graph theory is often combined with other related techniques to make up for the image generated by the unclear boundary.

2.4. Active Contour Models

The active contour model [14] can transform the process of image segmentation into the process of solving the energy functional minimum, extract the edge of the target, and define an energy functional to make its independent variable include the edge curve, so the process of segmentation becomes the process of solving the energy functional minimum. General energy functional segmentation methods include two categories: Parametric active profile model and geometric active profile model.

Energy functional segmentation method is more accurate in target location and segmentation precision, but the calculation is more complicated, and there may be isolated points and lines after processing.

3. Medical Image Segmentation Method based on Deep Learning

With the rise and development of big data, artificial intelligence and other theoretical technologies, the method based on deep learning provides a new idea for the realization of medical image segmentation. This part mainly introduces several classic deep learning image segmentation frameworks.

3.1. Fully Convolutional Networks

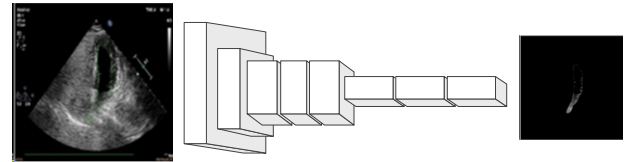


Figure 2. FCN structure

Fully Convolutional Networks (FCN), proposed by Long [15] et al in 2015, are composed of input, convolutional, pooling and output, the structure of which is shown in Figure 2. FCN uses Convolutional layers to replace the fully connected layers behind Convolutional Neural Networks (CNN), and implements pixel-level classification of images. It convolves the input image for several times to obtain the low-resolution feature information of the image, and then uses the last convolutional layer to up-sample the convolved image to restore it to the original size. Compared with CNN, FCN can input images of any size without repeated storage between pixel blocks, so segmentation efficiency is more efficient. However, it is not sensitive to the details of the target, segmentation is not accurate enough, and the correlation between pixels is not fully considered, resulting in the loss of some spatial information.

3.2. The Series of UNet

In 2015, Ronneberger [16] et al. proposed U-net network as a lightweight convolutional neural network with complete symmetry composed of convolution, downsampling, upsampling and skip-connection. It is capable of using a small amount of medical data. To obtain accurate segmentation results by integrating global and local detail features. In the

coding stage, feature extraction is carried out by double-layer convolution operation. Secondly, the dimensionality of the extracted features is reduced through downsampling, irrelevant feature information is removed, and the main feature information in the image is extracted after four dimensionality reduction. In the decoding stage, the feature after dimension reduction is up-sampled to restore the image to the original resolution. Then skip connection is used to introduce shallow information. Finally, the 1*1 convolution layer is used for classification, and the foreground and background layers are output, as shown in Figure 3.

In 2018, Zhou et al. [17] proposed UNet++, which replaced long connections in Unet network to integrate long connections and short connections, so that it could capture features of different levels. In 2020, Huang et al. [18] proposed UNet3+ on the basis of UNet and Unet++, which uses full-scale jump join and deep supervision to combine high-low level semantics from feature graphs of different scales and learn hierarchical representations from them, realizing full automation of segmentation.

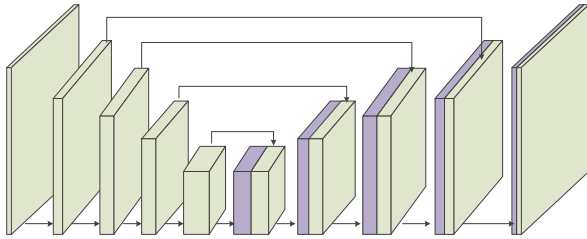


Figure 3. Unet structure

3.3. MedT

In the past decades, convolutional neural networks have been widely used in the field of medical image segmentation, and achieved good segmentation results in this field. However, CNN lacks understanding of remote dependency in images due to its inherent inductive bias. In recent years, transformer framework, which is widely used in natural language processing, has opened the door of image. MedT[19] published in 2021 puts forward a local-global training strategy, which not only operates the medical image as a whole, but also pays more attention to the local information of the image. Its performance is better than Res-UNet, U-net ++ and other networks, which further improves the accuracy of medical image segmentation.

Subsequently, many transformer frameworks are flooding into the field of medical image segmentation, such as MCTrans [20], UTrans[21] and CNS-Net [22], etc.

4. Summary

Medical image is an important reference for doctors to diagnose diseases. Effective segmentation of medical image is helpful for doctors to diagnose diseases, so as to improve diagnosis efficiency. This paper mainly introduces several traditional and deep learning-based medical image segmentation methods, but in the actual scene, different methods should be used for different images. Models based on deep learning can find hidden information from a large number of pictures without prior knowledge for feature extraction, which can achieve higher accuracy. However, most models are based on supervised learning, which requires a large number of segmentation labels, which wastes a lot of manpower and material resources to a certain extent. Therefore, finding better segmentation models and using less

labels is also a development direction of medical image segmentation in the future.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grants No. 62002102, and in part by Henan Postdoctoral Foundation, and Foundation from the Postdoctoral Research Station of Control Science and Engineering at Henan University of Science and Technology, and in part by the Key Technologies R & D Program of Henan Province under Grants No. 202102210169 and No. 212102210088.

References

- [1] Doi K. Computer-aided diagnosis in medical imaging: Historical review, current status and future potential [J]. *Computerized Medical Imaging and Graphics*, 2007, 31(4-5):198-211.
- [2] Wang D, Zhang Y, Zhang K, et al. Focalmix: Semi-supervised learning for 3d medical image detection[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020: 3951-3960.
- [3] Liu Y, Zhang F, Zhang Q, et al. Cross-view correspondence reasoning based on bipartite graph convolutional network for mammogram mass detection[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020: 3812-3822.
- [4] Wang Y, Wei X, Liu F, et al. Deep distance transform for tubular structure segmentation in ct scans[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020: 3833-3842.
- [5] Yu Q, Yang D, Roth H, et al. C2fnas: Coarse-to-fine neural architecture search for 3d medical image segmentation [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020: 4126-4135.
- [6] Wang J, Zhang M. Deepflash: An efficient network for learning-based medical image registration[C]//*Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020: 4444-4452.
- [7] Mok T C W, Chung A. Fast symmetric diffeomorphic image registration with convolutional neural networks [C]//*Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020: 4644-4653.
- [8] Al-Amri S S, Kalyankar N V. Image segmentation by using threshold techniques [J]. *arXiv preprint arXiv:1005.4020*, 2010.
- [9] Li F, Ng M K, Zeng T Y, et al. A multiphase image segmentation method based on fuzzy region competition[J]. *SIAM Journal on Imaging Sciences*, 2010, 3(3): 277-299.
- [10] Javadpour A, Mohammadi A. Improving brain magnetic resonance image (MRI) segmentation via a novel algorithm based on genetic and regional growth[J]. *Journal of biomedical physics & engineering*, 2016, 6(2): 95.
- [11] Chen P C, Pavlidis T. Segmentation by texture using a co-occurrence matrix and a split-and-merge algorithm[J]. *Computer graphics and image processing*, 1979, 10(2): 172-182.
- [12] Peng B, Zhang L, Zhang D. A survey of graph theoretical approaches to image segmentation[J]. *Pattern recognition*, 2013, 46(3): 1020-1038.
- [13] LaRocca F, Chiu S J, McNabb R P, et al. Robust automatic segmentation of corneal layer boundaries in SDOCT images using graph theory and dynamic programming[J]. *Biomedical optics express*, 2011, 2(6): 1524-1538.

- [14] Kass M, Witkin A, Terzopoulos D. Snakes: Active contour models [J]. International journal of computer vision, 1988, 1(4): 321-331.
- [15] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation [C]. Proceedings of the IEEE conference on computer vision and pattern recognition, 2015: 3431-3440.
- [16] Ronneberger O, Fischer P, Brox T. U-net: Convolutional Networks for Biomedical Image Segmentation [C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich: MICCAI, 2015: 234-241.
- [17] Zhou Z, Rahman Siddiquee M M, Tajbakhsh N, et al. Unet++: A nested u-net architecture for medical image segmentation [M]// Deep learning in medical image analysis and multimodal learning for clinical decision support. Springer, Cham, 2018: 3-11.
- [18] Huang H, Lin L, Tong R, et al. Unet 3+: A full-scale connected unet for medical image segmentation [C]// ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2020: 1055-1059.
- [19] Valanarasu J M J, Oza P, Hacıhaliloglu I, et al. Medical transformer: Gated axial-attention for medical image segmentation [C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2021: 36-46.
- [20] Ji Y, Zhang R, Wang H, et al. Multi-compound transformer for accurate biomedical image segmentation [C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2021: 326-336.
- [21] Gao Y, Zhou M, Metaxas D N. UTNet: a hybrid transformer architecture for medical image segmentation [C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2021: 61-71.
- [22] Ji G P, Chou Y C, Fan D P, et al. Progressively normalized self-attention network for video polyp segmentation [C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2021: 142-152.