Medical Image Segmentation Analysis and Research

Xiaohan Liu[a]
School of Integrated Circuit Science and Engineering, Tianjin University of Technology, TianJin, 300384, China

Keywords: Medical Image Segmentation, Region Segmentation, Deep Learning.

Abstract: Medical image segmentation technology, a vital part of medical imaging analysis, has made great progress in recent years. It is extremely important for the early identification of diseases, the making of treatment plans, and surgical planning. In the early days, traditional image segmentation methods, like those based on threshold-based segmentation, edge-detection, and region-growing, were effective in some simple scenarios. However, when they were faced with complex medical images, they often encountered challenges such as difficulty in handling noise interference, blurred boundaries, and multi-target overlapping. This paper first systematically reviews three traditional medical image segmentation techniques based on threshold, edge, and region, and then focuses on recent deep-learning-based segmentation techniques, including U-Net, Mask R-CNN, and DeepLab models. This paper also summarizes the current status of medical image segmentation techniques through examples of cell and organ segmentation as well as stomach cancer segmentation. Finally, from the aspects of deep learning model optimization and technology integration, this paper looks into the future of medical image segmentation technology.

1 INTRODUCTION

Medical image segmentation technology is a crucial element in medical image analysis. This holds substantial importance in the initial detection of diseases and the development of treatment plans during the early stages of disease onset. Its objective is to precisely extract target regions, such as organs, diseased tissues, or cellular structures, from complex medical images, thereby providing a reliable foundation for clinical decision-making. Given the significant differences in the segmentation of medical images by humans, previous researchers have conducted extensive research on medical image segmentation methods.

Traditional image segmentation methods, including threshold segmentation, edge detection and region growth segmentation methods, can perform basic image segmentation tasks in specific scenes through underlying features like image gray value, texture and spatial distribution. However, when dealing with complex medical images, traditional segmentation methods have obvious limitations. Image features are easily disturbed by noise, decreasing the segmentation accuracy. In multi - target scenarios, traditional image segmentation methods have difficulty in effectively differentiating targets from backgrounds, particularly when target and background features are alike, leading to subpar segmentation results. Moreover, when the target boundary in the image is unclear, traditional methods are likely to produce discontinuous or incorrect segmentation results, further undermining the segmentation accuracy. Since many traditional methods rely on manually setting seed points or parameters, this not only increases the operational complexity but also makes the segmentation results vulnerable to subjective influence.

In recent years, deep-learning technology has advanced rapidly, bringing revolutionary breakthroughs in medical image segmentation. By automatically learning multi-level features, deep-learning methods remarkably enhance the precision and stability of segmentation. This paper intends to explore the applications in medical image segmentation of U-Net, Mask R-CNN as well as DeepLab, along with their pros and cons, by conducting a contrastive analysis between traditional methods and deep-learning techniques. It provides

[a] https://orcid.org/0009-0001-2277-7078

Liu, X.
Medical Image Segmentation Analysis and Research.
DOI: 10.5220/0013680900004670

In *Proceedings of the 2nd International Conference on Data Science and Engineering (ICDSE 2025)*, pages 194-198
ISBN: 978-989-758-765-8

references for researchers in related fields and offers ideas for further research directions.

2 THRESHOLD SEGMENTATION

Threshold segmentation ranks among the simplest approaches in image segmentation. Its principle entails analyzing the grayscale values of an image, establishing a threshold, and comparing the pixels of the image with it. Then, distinct areas are separated according to the comparison results. Based on these, the components of the object and the backdrop within the picture can be recognized. A global threshold is selected by calculating the maximal inter-class variance within the image. The Otsu algorithm selects the threshold by calculating the minimum intra-class variance in addition to the maximum inter-class variance calculation (Otsu, 1975). Calculating the local threshold involves calculating the threshold in different areas of the image and then calculating the average threshold or dynamically selecting the threshold according to the alteration of pixel gray values. Ilhan U et al.. proposed a threshold-based method that segmented the tumor region from the brain MRI image by enhancing the brain MRI image (Ilhan, 2017). Maolood et al. proposed a segmentation method based on fuzzy entropy and level-set thresholds that can segment diverse cancer-related images (Maolood, 2018).

Threshold-based segmentation methods are simple in logic, fast in calculation, and can perform real-time processing. However, when handling complex structures, the contrast of gray values for the target area and the background area isn't evident, and they are prone to being interfered with by noise. At this time, the performance of threshold-based image segmentation will degrade, making it difficult to adapt to multimodal image features.

3 EDGE DETECTION

Medical image analysis based on edge detection aims to analyze the changes of gray value, texture, color and other features in the image. In an image, when an abrupt change occurs between pixels, the method uses an edge detection operator to detect the image target and tracks the contours to segment the image target. Commonly used edge detection operators include the Sobel operator (Sobel, 1978), the LoG operator (Marr and Hildreth, 1980), and the Canny operator (Canny,1986). The Sobel operator determines the edges by calculating the gradients in the horizontal and vertical directions, but it has difficulty in distinguishing the subject from the background strictly. The LoG operator is first smoothed by a Gaussian filter and then takes the second-order derivative to find the zero-crossing point. However, the operator is sensitive to noise, which leads to insufficient edge localization accuracy.

The Canny operator, based on a multi-stage algorithm, is more complex. It can effectively detect weak edges and suppress noise at the same time. However, the computation procedure is intricate and the processing time is extended. In practical image segmentation applications, these operators can be used to segment bone X-ray images (Lu, Tang, and Liu, 2023). blood vessel images, and lung CT images (Zhang, Fu, and Dai, 2019).

Segmentation based on edge detection is a method for segmenting images that have uniform regions. If the edge of the image is fuzzy or the image has more details, the accuracy of the image segmentation results will be low. In addition, since the method is based on the change of image grayscale, the obtained result is only the segmentation result of the image, which is not necessarily the same as the actual segmentation. At the same time, for some complex medical images, such as images of richly textured soft tissues, some important information may be lost.

4 REGIONAL GROWTH

To cluster pixels with similar characteristics together, researchers can use several methods in the division of medical images based on region growth. The region-growing approach selects several seed pixels on the picture. Subsequently, it combines the adjacent pixels with these seed pixels based on characteristics such as grayscale value, texture, and color, until no more pixels satisfying the criteria are available. This method relies on the setting of seed points and thresholds, and is likely to cause incomplete segmentation or over-segmentation. The region splitting and merging approach partitions the initial image into numerous regions. After that, it consecutively expands or combines these regions according to the similarity among the pixels. Once the features of the pixels in a region exceed a threshold set in advance, it means that there are different targets in the region and they need to be further segmented. This method is computationally complex and has poor noise immunity.

The segmentation method based on region growth can, on the one hand, establish a graph model using local information, and then segment the images of pulmonary arteries and veins (Jimenez-Carretero, Bermejo-Peláez, and Nardelli, 2019). On the other hand, this method can also accomplish the segmentation of a brain magnetic resonance image (MRI) by automatically choosing seed points along with a genetic algorithm (Javadpour and Mohammadi, 2016).

5 IMAGE SEGMENTATION BASED ON DEEP LEARNING

Due to the limitations of the above-mentioned medical image segmentation methods in handling complex images, researchers developed a deep-learning-based segmentation approach. This approach is intended to further enhance the image segmentation results, enabling better handling of complex-image segmentation tasks.

5.1 U-Net

U-Net, a deep-learning-based structure for image semantic partitioning, is of great significance in medical image segmentation and ranks among the most prevalently utilized image - segmentation models (Siddique, Paheding, and ElkinIn, 2021) order to achieve higher segmentation accuracy with a small number of training images, Ronneberger et al. developed a U-Net model suitable for biomedical image segmentation under the influence of the full convolutional network (FCN)(Ronneberger, Fischer, and Brox, 2015). The encoder-decoder model was proposed by Long et al(Long, Shelhamer, Darrell, 2015). U - Net's fundamental structure is composed of two components. One is the contraction path(encoder), which extracts high-level features from the image while reducing the amount of data by implementing downsampling on the feature map. The other part is the expansion path (decoder), which gradually recovers the extracted feature information to a resolution close to the original image size through upsampling. In addition, U-Net uses skip connections to directly connect the feature maps of the encoder and decoder, by combining high-level global features while preserving local features, thus effectively improving the accuracy of image segmentation. U-Net's unique U-shaped symmetric architecture extracts high-level semantic features with the help of contraction paths and preserves low-level detail features with the help of expansion paths, so as to achieve the organic integration of the two and effectively restore the image resolution. With this characteristic, U-Net enjoys a remarkable edge in medical image segmentation and has emerged as one of the crucial approaches in this domain.

U-Net can perform end-to-end training. This implies that during the training process, it can directly learn from the inputted medical images and generate the corresponding segmentation outcomes. This learning approach enables U-Net to adapt to different types of medical images and segmentation tasks, including neuronal structure segmentation, cell segmentation, heart segmentation (Antonelli, Reinke, Bakas, and Farahani, 2022), and lung CT image segmentation (Chen, 2023). Building on the evolution of the U-Net model, researchers have derived a diverse range of variants. For example, the 3D U-Net model is applicable to stereoscopic data segmentation (Çiçek, Abdulkadir, and Lienkamp, 2016). The Attention U-Net can focus on the segmentation of a particular thing and is not affected by background and other factors. U-Net++ is a further enhancement of U-Net. Besides cell segmentation, it can also segment organs like brain tumors (Micallef, Seychell, Bajada, 2021). U-Net3+ fuses features of different scales through full-scale jump connections, resulting in higher segmentation accuracy.

5.2 Mask R-CNN

Unlike semantic segmentation, instance segmentation adds labels to each target object after detecting the target type, and a representative model in instance segmentation is the Mask R-CNN model. Ren et al. 's Faster R-CNN model is an algorithm for target detection(Ren, He, Girshick, 2016). He et al 's Mask R-CNN model improves the Faster R-CNN model from two aspects (He, Gkioxari, Dollár, 2017): (1) To attain precise image segmentation, Mask R - CNN adds a mask branch in parallel to predict the region of interest (RoI). (2) In the Mask R - CNN model, the RoIPool layer is substituted with the RoLAlign layer since the rounding operation of the RoIPool layer can result in image positioning deviation. While performing object detection, pixel-level image segmentation is carried out, which allows Mask R-CNN to segment accurately in scenarios with blurred object boundaries and complex shapes.

When performing image segmentation with Mask R-CNN, the input image is initially pre-processed through clipping and other operations, and then the feature map is obtained through the neural convolutional network, multiple anchor points are set for the feature map and candidate RoI is obtained. Binary classification and bounding box regression are performed on the regional suggestion network (RPN) for these ROIs. The filtered RoI undergoes classification, bounding box regression, and mask generation through the RoIAlign operation.

For multi-organ segmentation, Shu et al. 's improved Mask R-CNN model accurately segmented seven major organs, including the heart, lung, liver, and kidney, in the same image(Shu, Nian, Yu, 2020). In the segmentation of bone and soft tissue images, Felfeliyan, B et al(Felfeliyan, Hareendranathan, Kuntze, 2022) improved the Mask R-CNN model by adjusting the loss function, enabling it to accurately segment multiple parts such as shoulder, elbow, and wrist in MSI MRI images. This not only achieves good results in the field of multi-organ segmentation but also has excellent performance in the segmentation of specific tissues and parts, further expanding the application range of the Mask R-CNN model in medical image segmentation.

Mask R-CNN can be applied to the prediction and segmentation of diseased areas in medical images and provides a solid foundation for subsequent algorithm improvement. For example, the 3D Mask R-CNN algorithm can improve the accuracy of brain tumor image segmentation(Jeong, Kahn, Liu, 2020).

5.3 DeepLab

When the DeepLab model processes an image, it first preprocesses the image to ensure consistent data distribution. Next, it extracts features from the image by means of a deep convolutional network, which is the core part of the DeepLab model. Subsequently, the DeepLab model introduces dilated convolution to expand the receptive field, improve the image-segmentation accuracy, and maintain the resolution simultaneously. After that, the model introduces Spatial Pyramid Pooling (SPP) to perform pooling operations on the feature maps, further enhancing its performance. Finally, the model optimizes the segmentation-target boundary according to the similarity between pixels by using the fully-connected conditional random field (CRF). After undergoing the above-mentioned processing, the DeepLab model outputs an image of the same size as the input image. Each pixel in the image is assigned a category label to indicate the semantic category it belongs to.

Overall, these techniques have enabled DeepLab to achieve remarkable results in semantic segmentation. This success has drawn numerous researchers to extend DeepLab's applications to diverse fields. Among them, the medical-image analysis field, due to its unique value and challenges, has become an important area that many researchers focus on and actively explore for applications.

DeepLab improves performance of semantic segmentation by integrating various techniques. In the field of medical image analysis, Wang et al. proposed a model based on the DeepLab v3+ network structure to precisely recognize gastric cancer images and segment the cancerous areas(Wang, Liu,2021). DeepLab v3+ network architecture optimizes the use of dilated convolution, enabling it to merge multi-scale features and improve segmentation capabilities for complex scenes. The model proposed based on this structure has significantly advanced the segmentation of gastric-cancer images and also promoted the growth of convolutional neural networks in the medical field.

6 CONCLUSIONS

Medical image segmentation technology, as an important computer - assisted diagnostic technology, has become a crucial and essential technology in medical image analysis and exerts a significant influence in the fields of disease diagnosis, treatment planning and medical research. Traditional segmentation methods, depending on image features and algorithm logic, can quickly complete the segmentation task when the image features are obvious or the scene is relatively simple. However, as the complexity of medical images grows, traditional segmentation methods are gradually failing to satisfy the requirements of modern medicine. Meanwhile, deep-learning-based segmentation technology can effectively compensate for the deficiencies of traditional methods when segmenting complex medical images.

Nowadays, in the realm of medical image segmentation, besides common segmentation methods, there are also multimodal image fusion, multiscale fusion and deep learning combined with traditional processing techniques. These methods can better obtain to capture complex information. Although the segmentation technology of medical images has made notable progress, it still faces some challenges. For example, deep learning models necessitate numerous training models, model robustness is insufficient, and the evaluation of image segmentation results varies from person to person.

In the future, medical image segmentation technology may develop in terms of ongoing optimization and refinement of deep learning models and combination with other technologies, so as to obtain more comprehensive pathological characteristics and expand different application scenarios. This will further improve the accuracy, robustness, and application value of medical image segmentation and provide strong support for precision medicine and personalized treatment.

REFERENCES

Antonelli, M., Reinke, A., Bakas, S., Farahani, K., Kopp-Schneider, A., Landman, B. A., ... & Cardoso, M. J., 2022. The medical segmentation decathlon. Nature Communications, 13(1), 4128.

Canny, J., 1986. A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, (6), 679-698.

Chen, Z., 2023. Medical Image Segmentation Based on U-Net. In Journal of Physics: Conference Series (Vol. 2547, No. 1, p. 012010). IOP Publishing.

Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O., 2016. 3D U-Net: learning dense volumetric segmentation from sparse annotation. In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19 (pp. 424-432). Springer International Publishing.

Felfeliyan, B., Hareendranathan, A., Kuntze, G., Jaremko, J. L., & Ronsky, J. L., 2022. Improved-Mask R-CNN: Towards an accurate generic MSK MRI instance segmentation platform (data from the Osteoarthritis Initiative). Computerized Medical Imaging and Graphics, 97, 102056.

He, K., Gkioxari, G., Dollár, P., & Girshick, R., 2017. Mask R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (pp. 2961-2969).

Ilhan, U., & Ilhan, A., 2017. Brain tumor segmentation based on a new threshold approach. Procedia Computer Science, 120, 580-587.

Javadpour, A., & Mohammadi, A., 2016. Improving brain magnetic resonance image (MRI) segmentation via a novel algorithm based on genetic and regional growth. Journal of Biomedical Physics & Engineering, 6(2), 95.

Jeong, J., Lei, Y., Kahn, S., Liu, T., Curran, W. J., Shu, H. K., ... & Yang, X., 2020. Brain tumor segmentation using 3D Mask R-CNN for dynamic susceptibility contrast enhanced perfusion imaging. Physics in Medicine & Biology, 65(18), 185009.

Jimenez-Carretero, D., Bermejo-Peláez, D., Nardelli, P., Fraga, P., Fraile, E., Estépar, R. S. J., & Ledesma-Carbayo, M. J., 2019. A graph-cut approach for pulmonary artery-vein segmentation in noncontrast CT images. Medical Image Analysis, 52, 144-159.

Long, J., Shelhamer, E., & Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 3431-3440).

Lu, F., Tang, C., Liu, T., Zhang, Z., & Li, L., 2023. Multi-attention segmentation networks combined with the Sobel operator for medical images. Sensors, 23(5), 2546.

Maolood, I. Y., Al-Salhi, Y. E. A., & Lu, S., 2018. Thresholding for medical image segmentation for cancer using fuzzy entropy with level set algorithm. Open Medicine, 13(1), 374-383.

Marr, D., & Hildreth, E., 1980. Theory of edge detection. Proceedings of the Royal Society of London. Series B. Biological Sciences, 207(1167), 187-217.

Micallef, N., Seychell, D., & Bajada, C. J., 2021. Exploring the U-Net++ model for automatic brain tumor segmentation. IEEE Access, 9, 125523-125539.

Otsu, N., 1975. A threshold selection method from gray-level histograms. Automatica, 11(285-296), 23-27.

Ren, S., He, K., Girshick, R., & Sun, J., 2016. Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(6), 1137-1149.

Ronneberger, O., Fischer, P., & Brox, T., 2015. U-Net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18 (pp. 234-241). Springer International Publishing.

Siddique, N., Paheding, S., Elkin, C. P., & Devabhaktuni, V., 2021. U-Net and its variants for medical image segmentation: A review of theory and applications. IEEE Access, 9, 82031-82057.

Shu, J. H., Nian, F. D., Yu, M. H., & Li, X., 2020. An improved Mask R-CNN model for multiorgan segmentation. Mathematical Problems in Engineering, 2020(1), 8351725.

Sobel, I., 1978. Neighborhood coding of binary images for fast contour following and general binary array processing. Computer Graphics and Image Processing, 8(1), 127-135.

Wang, J., & Liu, X., 2021. Medical image recognition and segmentation of pathological slices of gastric cancer based on Deeplab v3+ neural network. Computer Methods and Programs in Biomedicine, 207, 106210.

Zhang, Z., Fu, H., Dai, H., Shen, J., Pang, Y., & Shao, L., 2019. ET-Net: A generic edge-attention guidance network for medical image segmentation. In Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part I 22 (pp. 442-450). Springer International Publishing.