

Application of Image Segmentation Algorithms in Computer Vision

Yuhang Xu *

School of Computer and Software, University of Science and Technology Liaoning, Anshan Liaoning, 114000, China

* Corresponding author Email: xuy21955@gmail.com

Abstract: In the field of computer vision (CV), image segmentation technology, as a fundamental part, has a crucial impact on the accuracy of subsequent image processing tasks. Image segmentation is not only a crucial transitional step from image processing to image analysis, but also a hot and difficult research topic in the field of CV. Although significant progress has been made in the research of image segmentation algorithms, existing segmentation algorithms may still face challenges in certain specific scenarios due to the complexity and diversity of images, making it difficult to achieve ideal segmentation results. In recent years, the rapid development of deep learning (DL) technology has brought new breakthroughs to the field of image segmentation. DL models, especially Convolutional Neural Networks (CNNs), can capture semantic information of images more accurately by automatically learning feature representations in images, thereby achieving more precise image segmentation. This article delves into the research and application of image segmentation algorithms in CV, with a focus on the application of DL in the field of image segmentation. With the continuous development of advanced technologies such as DL, it is believed that image segmentation technology will play a greater role in more fields in the future.

Keywords: Computer Vision; Image Segmentation Algorithms; Research and Application.

1. Introduction

Image segmentation is one of the fundamental technologies in the field of CV, and it is also the most basic method for early processing of image data [1]. In CV tasks, image segmentation plays a crucial role in dividing images into multiple regions with similar properties, which facilitates subsequent tasks such as image analysis, object detection, and scene understanding [2]. High quality image segmentation results can provide more accurate information for other CV tasks, thereby improving the performance of the entire system [3]. Therefore, studying efficient and robust image segmentation methods has significant practical significance and application value. With the continuous development of CV technology, image segmentation algorithms have also undergone an evolutionary process from simple to complex, from single to diverse [4].

Traditional image segmentation methods mainly rely on low-level features such as intensity, texture, and color of the image for region division [5]. Among them, threshold based algorithms divide the pixel values of an image into different categories by setting one or more thresholds, thereby achieving image segmentation [6]. This method is simple and easy to implement, but its segmentation effect is often unsatisfactory for complex scenes and images under variable lighting conditions. Edge based algorithms extract edge information by detecting areas in the image where pixel values change dramatically, thereby achieving image segmentation. This method has good performance for images with obvious edges, but its performance will be greatly affected when facing edge blur or noise interference. In addition to traditional image segmentation methods, there are also some clustering based algorithms, such as K-means clustering, fuzzy C-means clustering, etc. These algorithms divide similar pixels into the same category by calculating the similarity between pixels, thereby achieving image segmentation. However, these methods may face problems

such as high computational complexity and low segmentation accuracy when dealing with large-scale images or complex scenes [7].

In recent years, with the rise of DL technology, image segmentation algorithms based on DL have gradually become a research hotspot. The DL model can capture the semantic information of images more accurately by automatically learning high-level feature representations in images, thereby achieving more precise image segmentation. Compared with traditional image segmentation methods, DL based image segmentation algorithms have higher segmentation accuracy and stronger robustness. At present, DL based image segmentation methods are mainly divided into three categories: image semantic segmentation, instance segmentation, and panoramic segmentation. Image semantic segmentation is one of the important applications of DL in the field of image segmentation. Its goal is to assign a predefined category label to each pixel in the image, thereby achieving pixel level image understanding. With the development of CNN, the performance of image semantic segmentation has been significantly improved. Many CNN based semantic segmentation models have been proposed, which further improve the accuracy and efficiency of image semantic segmentation by introducing different network structures and optimization strategies.

2. The Application of DL in Image Segmentation

As a fundamental task in the field of CV, image segmentation aims to divide images into multiple regions with similar properties, providing basic data for subsequent tasks such as object detection and scene understanding [8]. However, due to the complexity and correlation of images themselves, traditional image segmentation methods often have poor performance in dealing with ambiguity and uncertainty [9]. In recent years, with the rapid development

of DL technology, especially the widespread application of CNN, significant progress has been made in the application of DL in image segmentation [10]. Traditional image segmentation methods, such as threshold based segmentation and edge based segmentation, mainly rely on the underlying features of the image for segmentation. However, these methods often struggle to achieve ideal segmentation results when facing complex image scenes and variable lighting conditions. In addition, traditional methods often require preprocessing and post-processing of images, which is cumbersome and has limited effectiveness. DL, especially CNN, can more accurately capture the semantic information of images by automatically learning deep level feature representations.

In image segmentation tasks, DL models can learn the spatial relationships and contextual information between pixels, thereby achieving fine segmentation of images. The application of CNN in image segmentation is mainly reflected in two aspects: firstly, by constructing an end-to-end network structure, the segmentation results of the image are directly

output; The second is to utilize the feature learning ability of DL to extract deep level features of the image, providing strong support for subsequent segmentation tasks. Its structure is shown in Figure 1. In terms of end-to-end network structure, some classic models such as fully convolutional networks (FCN) and U-Net achieve direct mapping from feature maps to pixel level segmentation results by introducing deconvolution or upsampling operations. These models gradually integrate multi-scale features of the image through multiple downsampling and upsampling operations, thereby improving the accuracy and robustness of segmentation. In terms of feature learning, DL models learn deep level feature representations of images through a large amount of training data. These features not only contain low-level information such as color and texture of the image, but also high-level information such as the shape and structure of the object. By utilizing these features, DL models can better understand the semantic content of images, thereby achieving accurate segmentation of images.

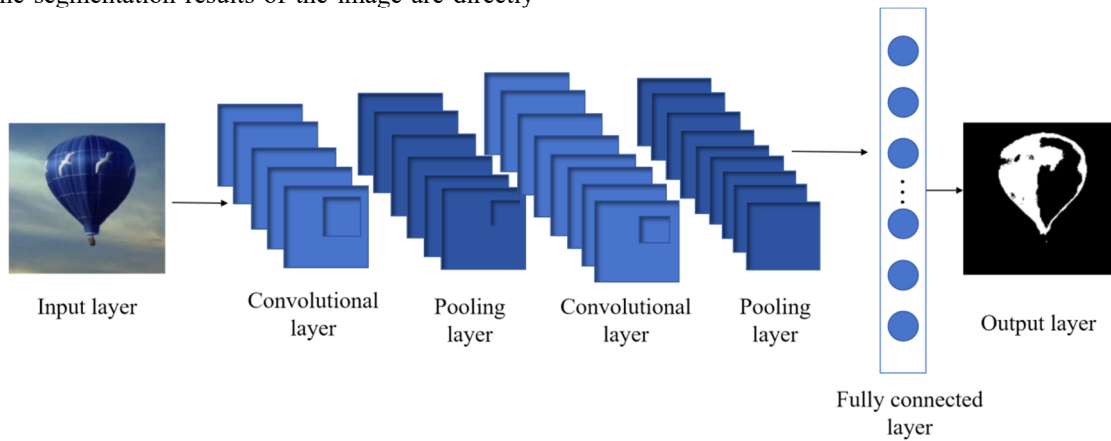


Figure 1. CNN structure

In addition, DL also combines some advanced optimization algorithms and loss functions to further improve the performance of image segmentation. For example, some models adopt a multi task learning approach that simultaneously learns image segmentation and classification tasks, improving segmentation accuracy by sharing underlying features. Some models also introduce attention mechanisms, allowing them to focus on key regions in the image, further improving segmentation accuracy. The application of DL in image segmentation not only improves the accuracy and efficiency of segmentation, but also provides more accurate data support for subsequent image processing and analysis tasks. For example, in the field of medical image analysis, DL models can achieve automatic segmentation and recognition of lesion areas, providing doctors with auxiliary diagnostic basis; In the field of autonomous driving, DL models can achieve precise segmentation of targets such as roads, vehicles, and pedestrians, providing assurance for the safe driving of vehicles.

3. Algorithm Principle

In the network, convolutional layers are the core components mainly used for feature extraction of images. Through convolutional operations, convolutional layers can learn local spatial patterns and texture information from the original image, and output corresponding feature maps.

However, traditional convolution operations often result in a loss of spatial resolution while increasing the receptive field, which may affect the accuracy of tasks such as image segmentation. To solve this problem, dilated convolution is introduced into the network. F_{out} represents the output image size after passing through the convolutional layer, F_{in} represents the input image size before passing through the convolutional layer, P represents *Padding*, k represents the convolutional kernel size, and s represents the convolutional kernel step size.

$$F_{out} = \left\lceil \frac{F_{in} + 2P - k}{s} \right\rceil + 1 \quad (1)$$

In CNN, the output of each convolutional layer is obtained by convolving the kernel with the input image, adding a bias term, and then applying an activation function. The size of the input image for each convolutional layer is $H \times W \times C$, where H and W represent the height and width of the image, C represents the number of channels, W_i is the weight of the i th convolutional kernel, b_i is the weight bias of the i th convolutional kernel, X_i is the input of the i th convolutional kernel, and f is the activation function. The formula for the output Y of the i th layer is:

$$Y = f(W_i \times X_i \times b_i) \quad (2)$$

In CNN, each layer of the network receives the output of the previous layer as its input. This hierarchical structure enables information to be transmitted and processed layer by layer in the network, ultimately resulting in a specific representation of input data. The activation function plays a crucial role in CNN. ReLU is one of the commonly used activation functions. The activation function ReLU is:

$$Y_i = \begin{cases} 0, & X_i < 0 \\ X_i, & X_i \geq 0 \end{cases} \quad (i = 1, \dots, N) \quad (3)$$

Among them, N is the size of the input feature map; X_i is the i th value of the input feature map; Y_i is the corresponding output.

4. Experimental Result

Experiment on image segmentation using MATLAB, using the Cityscapes dataset, which is a well-known dataset focused on semantic segmentation and scene understanding tasks in urban street scenes. This dataset is crucial for researching and

developing technologies in fields such as autonomous driving, robot navigation, and video surveillance. The Cityscapes dataset contains approximately 25000 images, which mainly capture street scenes in different cities, weather, and time conditions. The dataset defines 8 major categories (such as roads, sidewalks, buildings, vehicles, trees, etc.) and finer subclasses for more precise semantic segmentation tasks.

Figure 2 shows the comparison of image segmentation precision between different algorithms, which fully demonstrates the advantages of using DL technology in our algorithm. By introducing DL technology, our algorithm has achieved significantly better performance than traditional algorithms in image segmentation tasks. DL technology can automatically learn to extract useful features from raw data, which is a major advantage compared to traditional algorithms. By adopting a CNN structure, this algorithm can automatically learn hierarchical features in images. By constructing a deep neural network structure, algorithms can abstract and extract feature information from images layer by layer, thereby more accurately identifying different regions in the image. The ability of automatic feature learning enables DL algorithms to have higher flexibility and adaptability when dealing with complex image data.

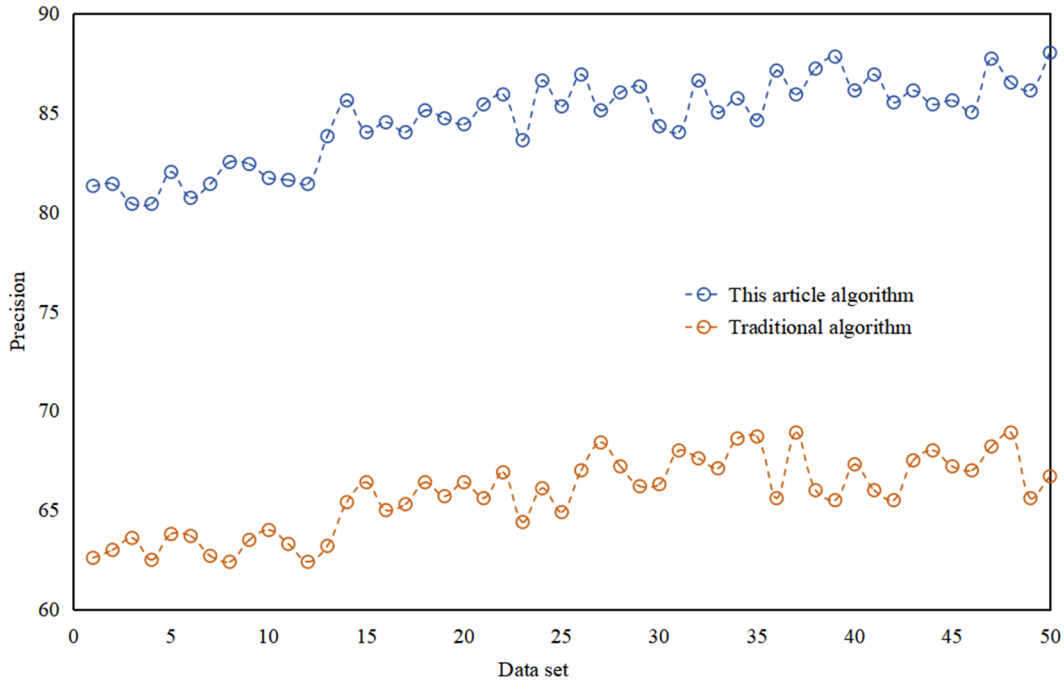


Figure 2. Comparison of segmentation precision

Figure 3 shows the loss function curve of the algorithm in this article. It can be seen from the figure that the loss function value gradually decreases with the increase of iteration times, indicating that the algorithm continuously learns and optimizes during the training process. Secondly, the speed of curve descent is relatively rapid in the initial stage, and then gradually stabilizes, which reflects the algorithm's ability to converge quickly in the early training stage and gradually reach a stable state in the later stage. This means that the algorithm proposed in this article can better cope with complex scenarios and changes, and has stronger robustness and adaptability.

5. Conclusion

With the continuous progress of computer technology and the increasing demand for image recognition precision by

humans, AI algorithms are gradually integrating into various industries, providing effective solutions for various problems. Especially in the field of image segmentation, the introduction of DL has brought unprecedented development opportunities in this field. Image segmentation, as one of the key technologies in CV, is of great significance for the understanding and analysis of images. With the rapid development of DL technology, especially the continuous improvement of models such as CNN, the performance of image segmentation algorithms has been significantly improved. These algorithms can automatically learn and extract complex features from images, achieving more accurate and efficient segmentation. With the widespread application of DL, image segmentation algorithms have become a research hotspot in the field of CV. This trend has driven the rapid development of image segmentation, providing strong support for many fields such as autonomous

driving, medical image analysis, and security monitoring. With the further development of computer technology and the

continuous enrichment of data resources, the performance of image segmentation algorithms will be further improved.

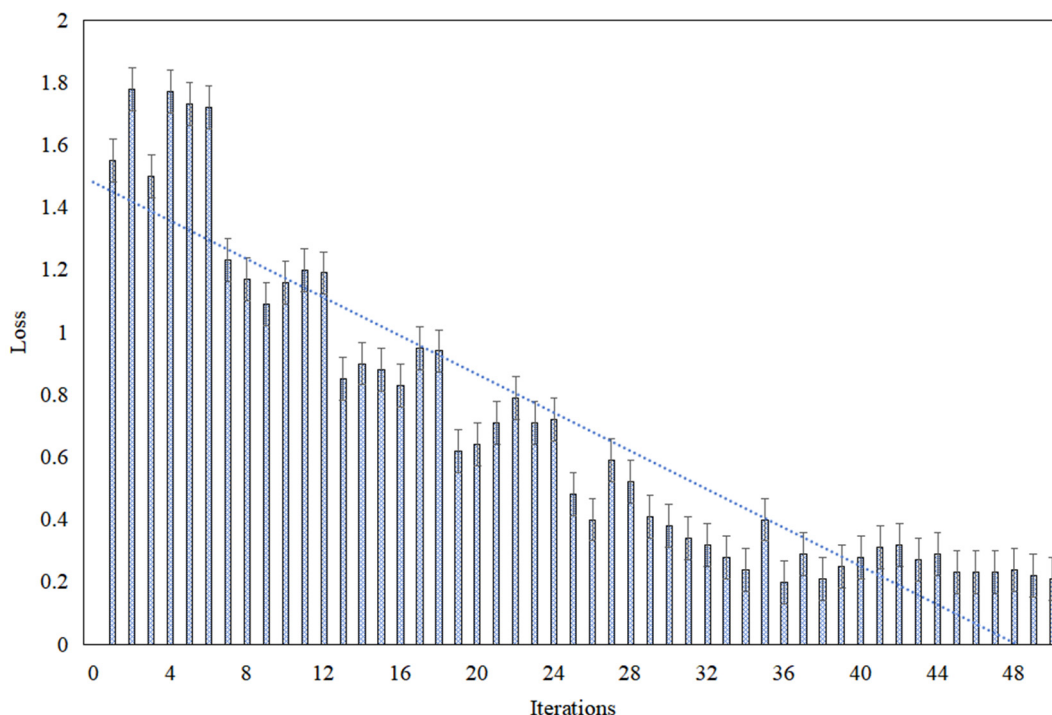


Figure 3. Loss function curve

References

- [1] Liu Siting, Han Xiaoxia, Zhao Xiaoning, et al. Image segmentation based on superpixel and density peak clustering [J]. Journal of Shanghai Electric Power University, 2023, 39 (2): 182-188.
- [2] Liu Yi, Zhang Xiaofeng, Sun Yujuan, et al. Robust image segmentation algorithm based on weighted filtering and kernel metric [J]. Progress of Laser and Optoelectronics, 2024, 61 (08): 09.
- [3] Tan Zhiguo, Ou Jianping, Zhang Jun, et al. Multi-feature combination depth image segmentation algorithm [J]. Computer Engineering and Science, 2018, 040(008):1429-1434.
- [4] Gu Pan, Zhang Fengdong. Weak supervised semantic image segmentation algorithm based on neural network [J]. computer applications and software, 2018, 35(2):5.
- [5] Li Xuan, Sun Xinnan. Image segmentation algorithm based on convolutional neural network [J]. Journal of Shenyang University of Aeronautics and Astronautics, 2020(1):50-57.
- [6] Guan Shenke, Lin Xiao, Zheng Xiaomei, Zhu Yuanyuan, Malizhuang. Multi-scale feature fusion image semantic segmentation algorithm combined with superpixel segmentation [J]. Acta Graphics, 2021, 042(003):406-413.
- [7] Zhang Xin, Zhang Xuyang, Mao Yuxin, et al. Overview of image semantic segmentation algorithms based on weak supervised learning [J]. Changjiang Information Communication, 2023, 36(5):105-108.
- [8] Lv Xin, Mu Xiaodong, Zhang Jun. Improved fuzzy C-means image segmentation algorithm based on spatial information [J]. Journal of Rocket Force University of Engineering (Natural Science Edition), 2020(2):8.
- [9] Chen Xingzhi, Wang Daiwen, Liu Naiyao, et al. Application of fast image segmentation algorithm model based on PSO-KMeans [J]. Modern Information Technology, 2020, 4(5):4.
- [10] Wu Shangzhi, Duan Chao, She Zhiyong. Image multi-threshold segmentation algorithm based on variable precision rough set and particle swarm optimization [J]. Journal of Northwest Normal University: Natural Science Edition, 2019, 55(2):8.