Overview of Image Segmentation

Lingdi Kong[a, *], Lingda Meng[b], Hongchen Li[c]

School of Software, Zhengzhou University, Zhengzhou 450000, China

[a]764452379@qq.com, [b]913896134@qq.com, [c]iot0lhc@vip.qq.com

*Corresponding author

Keywords: Image segmentation, Threshold segmentation, Neural Networks

Abstract: Image segmentation is the basis of image recognition and a difficult problem in computer vision research. As the first step of image analysis, image segmentation is to mark a similar part of the image or a part of interest. So far, although there is no universal image segmentation method, the algorithms developed at various stages represent the trend of image segmentation ideas, and are also suitable for image analysis in different occasions, from traditional image segmentation techniques to based on specific Algorithm segmentation method, and then the current semantic segmentation method based on neural network, here I will make a review of image segmentation algorithm.

1. Introduction

The research of image segmentation can be roughly divided into three parts. The first is the segmentation of similar parts based on pixels and traditional mathematics. Representative methods include threshold segmentation, region segmentation, edge and texture segmentation, etc. This type of method uses pixels as a unit and is suitable for simple images. The second part is the image segmentation method with the help of some mathematical models or tools. Typically, there are wavelet transform by frequency analysis and heuristic algorithms that use probability optimization such as genetic algorithm. The third part is semantic analysis based on neural networks and their variants. The most successful deep learning techniques for image segmentation are all based on a common pioneer: FCN (Fully Convolutional Network)[1], such kind of method combines more content-assisted analysis in the image to make up for the shortcomings of traditional methods that cannot segment internal complex segmentation blocks.

2. Traditional image segmentation method

The traditional method uses pixels or pixel areas as a unit to divide similar pixels or blocks into an area and then segment the original image. This method is simple and easy to use, suitable for image segmentation with simple tones, sharp contrast, and clear edges. Here are some typical algorithms:

2.1 Threshold segmentation method

This is the easiest method and easiest to think of. Set the threshold after converting the image to a grayscale image. The parts above or below the threshold are divided into different areas or segments. Multiple thresholds can be set if multiple regions need to be segmented.

There are two problems with threshold segmentation in practice: One is the selection of threshold. The other is the sensitivity to noise.


Figure 1. Examples of Traditional image segmentation method. [1]

2.2 Region segmentation method

Region segmentation can be divided into region growing algorithm and region splitting algorithm. The two methods can be understood as mutual inverse processes. Here is an introduction to the region growing algorithm. The region growing algorithm is to select some pixels in the seed area first, and then expand outwards according to certain rules, and classify the pixels that meet the conditions into a category until the no new pixels can be expanded.

The operation of region growth depends on the selection of seed pixels, expansion principle, and stopping conditions. [3] Area segmentation has improved threshold segmentation to a certain extent, and the customizable extension principle provides the possibility of image segmentation in complex situations and when the picture type is known.

2.3 Edge detection method

There must be a boundary between different regions, and the boundary between regions must be different from the area within the region. Finding this difference can find the region boundary and then segment the region. This is also one of the most studied algorithms. In order to find this difference, the derivative of the boundary area, the pixel area where the gray value changes sharply is the area boundary, which is the easiest solution to think of. In addition, it is also a method to transform the image to the frequency domain for analysis by Fourier transform.

2.4 Segmentation method based on graph theory

This kind of method refers to the knowledge of graph theory to image segmentation, and combines the previous methods, comprehensively considering the boundary, color, texture, gray scale and other aspects.

The common practice is to use each pixel as the vertex of the graph. The edges between the vertices connect each pixel. The weight of the edge represents the non-negative similarity of the connected pixels in terms of grayscale, color, and texture. [4] The cutting of the picture is to cut the graph. The principle during the segmentation process is to maximize the internal similarity of the divided regions after division, and the similarity of different regions is the smallest.

3. Segmentation method with specific theory

This type of method applies some theoretical tools to the image segmentation. Typically, frequency analysis methods based on wavelet analysis and intelligent optimization algorithm based on probability theory such as genetic algorithm.

3.1 Image segmentation method based on wavelet analysis

Wavelet analysis can be understood as an improved Fourier transform, which transforms the infinitely long trigonometric function basis in the Fourier transform into a finite-length continuously decaying wavelet basis. The overall idea is still to analyze the discontinuous grayscale change boundary of the image in the frequency domain.

Firstly, wavelet analysis can reduce the noise of the image, which can be combined with other methods to use the wavelet transform to reduce noise before input. Secondly, due to the drastic change of the gray scale of the regional boundary, wavelet transform can detect the boundary of the segmented region in the frequency domain and provide edge information. Wavelet analysis is essentially an edge detection method, so it also has its specific shortcomings, such as discontinuous edge points, etc., so in actual applications, it needs to be combined with other methods for analysis.

3.2 Image segmentation based on genetic algorithm

Genetic algorithm is a heuristic search optimization algorithm that mimics the evolution of biology. It first determines a population size as a random solution at the beginning, and then generates a new solution by simulating biological hybridization. The new solution is generated according to the established rules. Pursue the new solution generated according to the established rules to survive the fittest and enter the next group of random solutions. After many iterations, it converges to an approximate optimal solution. From this process, it can be seen that the genetic algorithm can perform concurrent search, and can try to avoid local optimal solutions.

To apply genetic algorithm to image segmentation, first you need to select the initial group. In this step, you need to consider the image background and target pixels representation methods by using chromosome and gene respectively; then determine the evaluation function, which reflects the advantages and disadvantages of segmentation; and finally design the convergence condition.

4. Semantic segmentation method based on neural network

Traditional methods have always tried to segment from the aspects of color, grayscale, texture, etc. However, these methods are only suitable for simple image segmentation in specific occasions, which is insufficient for segmentation in complex scenes and complex situations within regions. With the development of neural networks and the improvement of computing power, the deep learning technology combined with more information in the picture to assist image segmentation has occupied the mainstream status of image segmentation.

The semantic segmentation based on neural network method can be divided into the following stages: neural network research stage, convolutional neural network research stage, FCN fully convolutional neural network stage, FCN-based deep learning technology stage.

Neural network is a simulation of the human brain nervous system. At first, it was composed of a single layer of "perceptron", which could only do simple linear classification tasks. Later, with the improvement of computing power, a two-layer neural network (multi-layer perceptron) appeared. The two-layer neural network has been theoretically proved that it can approximate any continuous function, that is, it can perform complex nonlinear classification tasks. The subsequent neural networks have been applied in various fields, and a variety of neural networks have been extended, including deeper deep neural networks, CNN convolutional neural networks, and RNN recurrent neural networks. Among them, CNN convolutional neural network is used for feature extraction and image classification. The appearance of CNN is mainly to solve the large number of neurons that are almost impossible to calculate when the neural network is fully connected. CNN has three characteristics: local perception using convolution kernel, weight sharing, pooling. reference four Local perception introduces a convolution kernel to the original image to make the original full connection become local connection. Weight sharing refers to sharing the convolution kernel, which further reduces the amount of calculation. The pooling operation can be simply understood as compress the original image, which is used to reduce the size of the image, commonly used average or maximum. On the basis of CNN, the pioneer of the most successful deep learning technology in the field of image segmentation appears: FCN fully convolutional neural network.

4.1 Fully Convolutional Neural Network FCN

The basic idea of FCN is to extend image-level classification to pixel-level classification to achieve semantic image segmentation.

The so-called full convolution means that the fully connected layer in the CNN classification network is replaced with convolutional layer, so that the entire network is connected by convolution.

FCN will eventually output a segmentation map. After the convolution, the resolution of the image is reduced. Then, the upsampling process needs to be performed to enlarge the resolution to restore it to the same size as the input image. Finally, output the image. The upsampling process can be understood as the inverse process of pooling operation, which can also be called deconvolution. The whole process can be visually represented by the following picture:

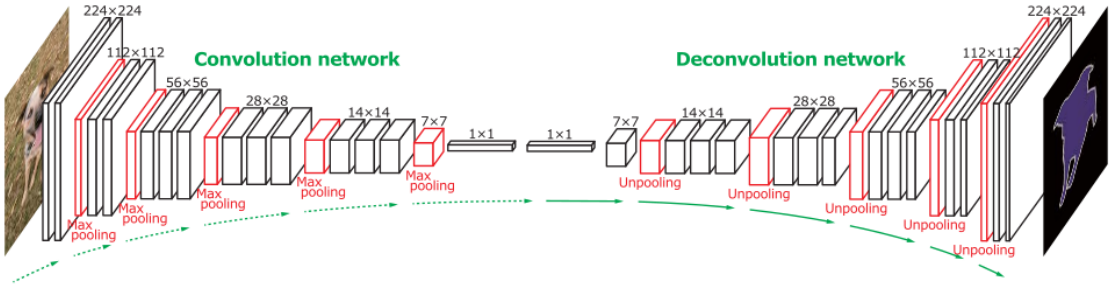

Figure 2. The process of deconvolution. [5]

The advantage of FCN is that it can accept any size of input, because there is no fully connected layer, and the operation efficiency is improved at the same time. The disadvantage is that the pixel-by-pixel classification does not consider the relationship between pixels and lacks overall consistency. Besides, some details will be ignored due to sampling.

4.2 SegNet semantic segmentation method

SegNet is an extension method based on FCN. Like FCN, SegNet removes the fully connected layer. SegNet consists of Encoder encoder and Decoder decoder to form an encoding-decoding structure, following a pixel classification layer. Among them, the Encoder encoder performs convolution and maximum pooling operations, and records the index position where the maximum value is located. In the Decoder process, the index is used to achieve nonlinear upsampling, so that there is no need to learn in the upsampling stage. Repeat the upsampling, and finally use the activation function to get the classification result.

Compared with FCN, SegNet has two advantages. First, SegNet does not need to save the entire feature map, but only saves the pooled index, which can save memory space. Second, SegNet does not use deconvolution, so there is no need to learn in the upsampling stage. In short SegNet reduces memory and time consumption.

5. Summary

This article summarizes the methods and ideas commonly used in the field of image segmentation. First, the traditional segmentation methods considered in terms of grayscale, color, texture, etc., followed by the application of some theoretical tool models to image segmentation, mainly listing wavelet analysis And genetic algorithm applications, and then introduces the current mainstream algorithms in the field of image segmentation-semantic segmentation models based on deep learning, from neural networks to convolutional neural networks CNN, to fully convolutional neural networks FCN, and finally introduces Algorithm SegNet based on FCN, which optimizes memory consumption and time consumption relative to FCN.

There is no specific method for image segmentation, and it is necessary to select an appropriate method in different situations. For some simple image segmentation occasions, filtering and threshold segmentation algorithms are sufficient. At the same time, we can also see that semantic-based image segmentation is more suitable for real life scenes. Image segmentation is still a developing technology. FCN-based extension methods continue to emerge. Innovative technologies for image segmentation should not only focus on the accuracy of segmentation, but also consider the real memory and time consumption issues.

References

[1] Alberto G.G, Sergio O.E, Sergiu O, Victor V.M, Jose G.R. (2017) A Review on Deep Learning Techniques Applied to Semantic Segmentation. ArXiv: 1704. 06857, 4: 8-13.

[2] Blowing Bubbles on Everest. (2018) Automatic threshold segmentation of images. https://www.cnblogs.com/Terrypython/p/9958250.html.

[3] Haoze Jing. (2019) the most complete overview of image segmentation. https://www.cnblogs.com/CV-life/p/11160796.html.

[4] SIGAI. (2019) Image segmentation technology introduction.https://zhuanlan.zhihu.com/p/49512872.

[5] Linolzhang. (2017) Image segmentation and FCN.https://blog.csdn.net/linolzhang/article/details/71698570?depth_1utm_source=distribute.pc_relevant.none-task&utm_source=distribute.pc_relevant.none-task.