Evolution Analysis of Adaptive Distance-Preserving Level Set for Image Segmentation Based on Depth Learning

Jinpeng Wang a, Jing Li

College of Electronic and Information Engineering, Shandong University of Science and Technology, Qingdao 266590, People’s Republic of China.

a 1561829762@qq.com

Abstract. Image segmentation technology is the most basic part of computer vision and the basis of all other image processing methods. The quality of image segmentation technology will affect the effect of subsequent processing to a large extent. In the research and application of computer vision, image segmentation is often the first step and plays a very important role in the whole process. Compared with image classification and target detection, Distance-Preserving level set method involves two cross-cutting areas of image recognition and natural language processing, which is very challenging. Handling the information interaction between the machine and the real world is the first step in the study of artificial intelligence. The goal of computer vision research is to make the computer have the ability to perceive and understand image information. In the future, the focus of the work will be on improving the generalization ability of the model so that the model can perform well in a variety of tasks.

1. Introduction

Computer-generated image description is a hotspot and difficulty in the current computer research field. Its basic task is to make the computer correctly describe the perceived image scene and the content of the scene [1]. In the digital information age, people have to accept a large amount of information every day, most of which are image information obtained by human vision, so image is an important information carrier for human perception of the world [2]. Image segmentation is an important and low-level processing task in computer vision. The quality of segmentation has a great impact on the completion of subsequent tasks. In the research and application of computer vision, image segmentation is often the first step, and plays a very important role in the whole process [3]. Faced with the huge amount of data that is constantly updated, human understanding can neither complete the perception of large amounts of data, and its understanding speed cannot match the update speed of information [4]. Therefore, it is easy to cause problems such as a series of key targets and situational perception errors or untimely perceptions. The computer processing image description task not only focuses on the object recognition in the image, but also pays attention to the mutual connection between the objects in the image [5].

Handling the information interaction between the machine and the real world is the first step in the study of artificial intelligence. The goal of computer vision research is to make the computer have the ability to perceive and understand image information similar to humans [6]. Since there is no image segmentation theory that can be widely used, there is no evaluation criterion for segmentation results suitable for all segmentation scenarios. Therefore, most of the various segmentation algorithms that have been proposed so far are specific to a specific problem [7]. There is no universal image segmentation algorithm that can fit all application scenarios. Compared with image classification and target detection,
the distance-maintaining level set method involves two major cross-fields of image recognition and natural language processing, which is extremely challenging [8]. By introducing weight coefficients related to image information, the size can be adjusted adaptively according to image information. Thus, the speed of curve evolution is accelerated, and the segmentation effect of the algorithm is improved [9]. Using the gray mean difference of small neighborhoods inside and outside the contour of evolutionary curve as an independent variable, an adaptive weight function model is established, which can improve the efficiency and accuracy of segmentation of gray-scale inhomogeneous images.

2. Image Segmentation Based on Deep Learning

2.1. Structure Analysis of Deep Neural Network
The distance-maintaining set method adds an internal energy functional to the energy functional of the traditional method, which is mainly used to correct the deviation of the level set function from the symbol distance function. The image is represented by a vector, and the decoding end uses a cyclic neural network language model for generating a language description. In the case of a large number of input and output mapping samples, the machine can be used to find a way to simulate this mapping mode [10]. Because such mapping is not quantifiable and difficult to prove. You can only move closer to the real map as much as possible. The neurons in each layer of the neural network are fully connected with each other. It is feasible to use this fully connected network to deal with handwritten numeral recognition, but the recognition effect is poor. The main reason is that for image data, there is a correlation between adjacent pixels of the image. In the traditional geometric active contour model, the level set function must always be approximately symbolic distance function in the evolution process, so as to ensure the stability of the level set evolution and the validity of the results.

The gradient of the image varies greatly in the noise area, which causes the variable weight coefficients to be anonymous on both sides of the noise. The model algorithm falls into the local minimum here, which makes the model more sensitive to the noise. For the task of target classification recognition, the quality of feature extraction directly affects the accuracy level of recognition. The uneven gray level image may have a large gradient value at the non-edge points, which results in variable weight coefficients having different sign on both sides of the image. And the edge stopping function is close to zero, so the segmentation of uneven gray image is not ideal. However, the high time complexity and a large number of redundant windows seriously restrict the efficiency of target location awareness.

2.2. Image Target Category Perception
From the perspective of visual perception, image segmentation seems to be a very easy problem to solve. In actual image processing, it is a very difficult task. The simplest and most common way to reduce image data over-fitting is to use label-preserving transformation method to artificially expand training data sets. The edge stopping function may reach the local minimum at the noise point, which makes the model more sensitive to noise. Unsupervised training is used to pre-train the network, and the weights are initialized with the results of pre-training. Then the parameters are fine-tuned through the supervised training team network model. For image data, one way to extend the data set is to pan and flip the image without changing the image tag. For some non-edge parts due to the uneven gray scale of the image, a large gradient value is generated, and the edge stop function approaches zero here, making the model inaccurate for such image segmentation. The distance-maintaining set method lacks the adaptability of direction and size. The weight coefficient not only determines the evolution direction of the zero level set, but also changes the amplitude of the stop speed function.

After extracting the target candidate regions, it is also necessary to perform feature extraction on these regions and then complete the category identification of the target. For a certain target in the image scene, all the candidate regions including the target are sorted according to the category score from large to small, and the candidate region with the highest score is unconditionally retained. Then, the ratio of the overlapping area of each of the remaining areas to the area and the smaller area of the two areas is calculated, and if the ratio exceeds the set value, the area with the lower score is suppressed. Since the
different types of targets are different in size in the image, the position of occurrence is also uncertain. It is necessary to design sliding windows of different scales and sizes to repeatedly detect images in order to complete the position perception of all categories of targets.

3. Evolution of Distance-Preserving level set

The combination of in-depth learning network and a large number of training data makes it achieve good results in various fields, and far away from other machine learning algorithms. Initialization curves must surround or inside the target object, otherwise wrong boundaries will occur or no boundaries can be extracted at all. The model based on distance regularization and gradient information solves the problem that the traditional geometric contour model needs to be reinitialized continuously. Referring to the spatial characteristics of the pixels in the image, each pixel is only closely related to the surrounding pixels in a certain range. Convolutional neural network is designed as a sparse network, that is, each neuron is only connected with the adjacent neurons. The model may reach local minimum values at the image noise and the non-boundary gradient, which makes the model's noise resistance poor and inaccurate for grayscale uneven image segmentation.

For some images, some parts of the target object are beyond the boundaries of the image. No matter how the initial curve is placed, it cannot meet the requirements of surrounding the target object or being inside or outside the target object. The performance parameters of image segmentation before and after optimization are shown in Table 1. After wavelet variation and filtering optimization, the image segmentation topology is greatly optimized, with fewer nodes and better monitoring area. Image segmentation topology reliability optimization simulation comparison is shown in Figure 1.

| Table 1. Performance parameters of image segmentation structure before and after optimization |
|---------------------------------|-----------------|-----------------|
| Before optimization          | After optimization |
| Row number                     | 136             | 112             |
| Column number                  | 97              | 65              |
| Monitoring points              | 346             | 534             |

Figure 1. Simulation comparison of image segmentation filtering optimization

With the rapid development of smart devices and the Internet, the collection and acquisition of common scene images has become simple and convenient. In the adaptive distance-maintaining level set evolution method, the weight coefficients are the same on both sides of the target boundary, making the model lack directionality. The obtained feature maps are fused according to the specified step size. The fusion method used here is additive. In the image obtained by fusion, the relative positions between the feature maps are the same as the relative positions between the original pixel points. In a complex image of a scene, there are many targets involved. Different kinds of targets occupy different proportions of pixels in the same image. Even for the same type of target, the ratio of image pixels to image pixels will vary. Subject objects often occupy most of the pixels of the image. Generally speaking, the
distribution of such objects in the image is not very dense. Researchers in the field of computer vision can freely obtain a large number of visible scene images through the network. Although this brings convenience to the training and testing of the model, because the data collection largely depends on the subjectivity of researchers, the test results are also lack of comparability. As far as radar image is concerned, whether it is the SAR needed for acquisition or the graphical restoration of radar data for special research, it is very professional.

An input does not affect only one output, but all subsequent outputs. For the extraction of target candidate regions, the initial method is to use sliding window to traverse the image, cut different regions of the image for feature extraction, and then identify whether there is a target in the region. The adaptive Distance-Preserving level set evolutionary model replaces the constant weight coefficients of the original model with variable weight coefficients, which can get rid of the dependence of the evolutionary curve on the initial position. The output of image segmentation is an end-to-end pixel-level classification label, requiring the output to have the same size as the input. After multiple convolutions and pooling operations in the encoding phase, the output size is getting smaller and smaller. Most of the input image is adapted to a fixed size by cropping or zooming, but the cropped area does not necessarily contain the target. Zooming will result in geometric distortion of the target, and the recognition accuracy will be reduced due to missing or deformed content. It is necessary to find a reasonable target candidate region extraction method, and reduce the number of target candidate regions, thereby reducing the calculation amount of the entire target location sensing system.

4. Conclusion

Deep learning based algorithms have become the mainstream method in the field of image segmentation, and various novel and efficient network models are constantly being proposed. The distance maintaining level set method is a major breakthrough in the traditional level set method. It does not need to periodically initialize the level set function, and in the numerical calculation of the evolution equation, a simple difference format can be used and a larger time step is allowed. The decoding phase uses transposed convolution to restore features to their original size step by step. At the same time, the output fusion corresponding to the coding phase further utilizes the intermediate results. The model not only has high robustness to noise and initial contour, but also can segment gray scale inhomogeneous image quickly and accurately. Deep learning technology can be used in situational element detection tasks in situational awareness system, and its perception accuracy and efficiency have reached a high level. It effectively solves the problem that the evolutionary curve is sensitive to the initial position. It can define the initial curve in any direction. The zero level set can decide whether to move inward or outward adaptively according to the nature of the image. In the future, we will focus on improving the generalization ability of the model, so that the model can perform well in a variety of tasks.

References
[1] Min H, Jia W, Wang X F, et al. An Intensity-Texture model based level set method for image segmentation[J]. Pattern Recognition, 2015, 48(4):1547-1562.
[2] Zhang K, Liu Q, Song H, et al. A Variational Approach to Simultaneous Image Segmentation and Bias Correction[J]. IEEE Transactions on Cybernetics, 2015, 45(8):1426-1437.
[3] Hu M H, Dong Q L, Liu B L, et al. The Potential of Double K-Means Clustering for Banana Image Segmentation[J]. Journal of Food Process Engineering, 2014, 37(1):10-18.
[4] Al-Ayyoub M, Abu-Dalo A M, Jararweh Y, et al. A GPU-based implementations of the fuzzy C-means algorithms for medical image segmentation[J]. The Journal of Supercomputing, 2015, 71(8):3149-3162.
[5] Zhang H, Ye X, Chen Y. An Efficient Algorithm for Multiphase Image Segmentation With Intensity Bias Correction[J]. IEEE Transactions on Image Processing, 2013, 22(10):3842-3851.
[6] Zhou D, Zhou H. A modified strategy of fuzzy clustering algorithm for image segmentation[J]. Soft Computing, 2015, 19(11):3261-3272.
[7] Yaacobson F S, Givoli D. An adaptive finite element procedure for the image segmentation problem.[J]. International Journal for Numerical Methods in Biomedical Engineering, 2015, 14(7):621-632.

[8] Wang X, Guo L, Yin J, et al. Narrowband Chan-Vese model of sonar image segmentation: A adaptive ladder initialization approach[J]. Applied Acoustics, 2016, 113:238-254.

[9] Ayech M W, Ziou D. Terahertz image segmentation using k-means clustering based on weighted feature learning and random pixel sampling[J]. Neurocomputing, 2016, 175:243-264.

[10] Cho H, Kang S J, Cho S I, et al. Image segmentation using linked mean-shift vectors and its implementation on GPU[J]. IEEE Transactions on Consumer Electronics, 2014, 60(4):719-727.