Original Research Article

A two-stage strategy image segmentation algorithm based on deep plug and play

Qin Yaping, Liu Hongjian

Changsha Normal University, Changsha, Hunan, 410100, China

Abstract: Image segmentation is a fundamental and challenging task in the field of image processing. This article proposes a two-stage strategy image segmentation algorithm based on deep plug and play. By conducting numerical experiments on pyCharm and MATLAB, we have verified the robustness and superiority of our proposed algorithm.

Keywords: Image segmentation; Two-stage strategy; Deep plug and play

1. Introduction

Image segmentation is a fundamental and challenging task in the field of image processing. Image segmentation is the process of removing unwanted background information from an image and extracting valuable foreground targets. Image segmentation plays an important role in many application fields, such as medical imaging, autonomous driving, license plate recognition, and so on. The two-stage strategy image segmentation algorithm[1] decomposes the image segmentation task into two stages. The first stage obtains a smooth approximate solution by minimizing the functional model, and the second stage performs threshold processing on the obtained approximate solution to complete the segmentation. The segmentation algorithm is easy to conduct numerical experiments due to the convexity of its functional model, and the experimental results show that the model has certain effectiveness and superiority. In recent years, deep learning methods have become popular in the field of image processing due to their high learning efficiency and strong adaptability. This article attempts to improve the two-stage strategy image segmentation algorithm by introducing a deep learning module and proposing a two-stage strategy image segmentation algorithm based on depth plug and play to further improve image segmentation efficiency.

2. Proposal of a two-stage strategy image segmentation algorithm based on deep plug and play

We solve the following model in the first stage:

$$\min_u E(u) = \min_u \frac{\lambda}{2}\|f - \mathcal{A}u\|_2^2 + \frac{\mu}{2}\|\nabla u\|_2^2 + \Phi(u)\,, \tag{1}$$

Among them λ, μ are positive parameters, f is a known degraded image, u is an ideal image that needs to be solved, Φ is an implicit regularization operator, ∇ is a gradient operator, $\mathcal{A}$ is a linear operator related to the problem. When $\mathcal{A} = \mathcal{I}$, f is a noisy image; when $\mathcal{A}$ is the blur operator, f is the blurred image with noise. In the second stage, we will perform threshold processing on the solution u from the first stage to complete the segmentation..

3. Solution of two-stage strategy image segmentation algorithm based on deep plug and play

Let $u=\left(u_1,u_2,...,u_{n^2}\right)^T$, and let u expand according to periodic boundary conditions, using first-order finite difference to approximate the discrete gradient operator ∇,

$$\|\nabla u\|_2^2=\sum_{i=1}^{n^2}\left(\left(\nabla^{(1)}u\right)_i^2+\left(\nabla^{(2)}u\right)_i^2\right),$$

where

$$\begin{cases}\left(\nabla^{(1)}u\right)_i=\begin{cases}u_{i+n}-u_i, & \text{if } 1\le i\le n(n-1),\\ u_{\mathrm{mod}(i,n)}-u_i, & \text{other.}\end{cases}\\ \left(\nabla^{(2)}u\right)_i=\begin{cases}u_{i+1}-u_i, & \text{if } \mathrm{mod}(i,n)\ne 0,\\ u_{i-n+1}-u_i, & \text{other.}\end{cases}\end{cases}$$

3.1. First phase

The first stage is to solve the optimization problem(1) .Firstly, introduce two auxiliary variables, z and w, and transform optimization problem (1) into the following problem:

$$\begin{aligned}&\min_u \frac{\lambda}{2}\|f-\mathcal{A}u\|_2^2+\frac{\mu}{2}\|\nabla z\|_2^2+\Phi(w)\\ &s.t.\ u=z, z=w\end{aligned}. \tag{2}$$

Then, using the augmented Lagrangian method, problem(2) is rewritten as:

$$\mathcal{L}(u,z,w)=\frac{\lambda}{2}\|f-\mathcal{A}u\|_2^2+\frac{\mu}{2}\|\nabla z\|_2^2+\Phi(w)+\frac{\beta_1}{2}\|u-z\|_2^2+\frac{\beta_2}{2}\|z-w\|_2^2, \tag{3}$$

Among them, β_1、β_2 is the penalty parameter. Problem(3) can be iteratively solved using the alternating direction multiplier method and can be decomposed into the following subproblems concerning u, z, and w:

$$\begin{cases}u_k=\arg\min_u \frac{\lambda}{2}\|f-\mathcal{A}u\|_2^2+\frac{\beta_1}{2}\|u-z_{k-1}\|_2^2\\ z_k=\arg\min_z \frac{\mu}{2}\|\nabla z\|_2^2+\frac{\beta_1}{2}\|u_k-z\|_2^2+\frac{\beta_2}{2}\|z-w_{k-1}\|_2^2\\ w_k=\arg\min_z \Phi(w)+\frac{\beta_2}{2}\|z_k-w\|_2^2\end{cases}. \tag{4}$$

(1) U-subproblem

The optimization conditions for the u -subproblem are satisfied

$$(\lambda\mathcal{A}^*\mathcal{A}+\beta_1)u=\lambda\mathcal{A}^*f+\beta_1 z_{k-1}. \tag{5}$$

Because $\mathcal{A}^T\mathcal{A}$ it is a block cyclic matrix, under the assumption that u is a periodic boundary condition, it can be solved using fast Fourier transform. The formula for solving u is:

$$u_k = \mathcal{F}^{-1}\left(\frac{\lambda\mathcal{F}(\mathcal{A}^* f) + \beta_1\mathcal{F}(z_{k-1})}{\lambda\mathcal{F}(\mathcal{A}^*\mathcal{A}) + \beta_1}\right). \tag{6}$$

In order to better describe the Fourier transform solving process, we can describe the degradation process of the image as

$f_a = u_a \otimes k + n_a$,

The subscript a represents the transformation of the n^2 dimensional vector into a $n \times n$ matrix,i.e. $u_a = array(u)$, $u_a \otimes k$ represents a clean image u_a and Blurring Kernel $k \in R^{m\times m}$ two dimensional convolution between them.when k is the delta(identity) kernel, f_a is the noisy image;when k is other non negative blur kernels, f_a is a blurred image with noise.Using the properties of Fourier transform and convolution, equation(6) can be written as:

$$u_k = vec\left(\mathcal{F}^{-1}\left(\frac{\lambda conj(\mathcal{F}(\kappa)).*\mathcal{F}\left(f_a\right) + \beta_1\mathcal{F}\left(\left(z_{k-1}\right)_a\right)}{\lambda\left|\mathcal{F}(\kappa)\right|^2 + \beta_1}\right)\right), \tag{7}$$

Among them, $\mathcal{F}$ represents the two-dimensional fast Fourier transform operator, $\mathcal{F}^{-1}$ represents the two-dimensional inverse fast Fourier transform operator, vec is the inverse of an $array$,representing the transformation of the $n \times n$ matrix into n^2 dimensional vectors by column. κ indicates that first k in the matrix of size $n \times n$ with zeros on the right and bottom sides.Then, with the middle row and column of k as the center, cycle the $n \times n$ matrix left and up by rolling the $\left[\frac{m}{2}\right]$ column and $\left[\frac{m}{2}\right]$ row respectively.

(2) Z-subproblem

The optimization conditions for z-subproblems are satisfied

$(\mu\nabla^T\nabla + \beta_1 + \beta_2)z = \beta_1 u_k + \beta_2 w_{k-1}$,

Since $\nabla^T\nabla$ is a block cyclic matrix,under the assumption of periodic boundaries, the Fast Fourier Transform can be used to solve:

$$z_k = vec\left(\mathcal{F}^{-1}\left(\frac{\beta_1\mathcal{F}\left(\left(u_k\right)_a\right) + \beta_2\mathcal{F}\left(\left(w_{k-1}\right)_a\right)}{\mu\mathcal{F}(t) + \beta_1 + \beta_2}\right)\right), \tag{8}$$

Among them $t = \begin{bmatrix} 4 & -1 & 0 & \cdots & 0 & -1 \\ -1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \in R^{n\times n}$.

(3)*W-subproblem*

The w-subproblem corresponds to a Gaussian denoising problem with a noise level of $\frac{1}{\sqrt{\beta_2}}$ on z_k .The w-sub problem can be rewritten as

$$w_k = denoiser\left(z_k, \frac{1}{\sqrt{\beta_2}}\right), \tag{9}$$

Denoiser is a Gaussian denoising algorithm based on deep learning. This article uses the DRUNet denoising algorithm[2]. The denoising structure combines U-Net and ResNet, achieving advanced denoising effects. The specific structure and the code can be found in[2].

3.2. Second phase

In the second stage, the first stage result u is segmented into K different regions to display image features. We use the MATLAB "KMEANS" command to achieve automatic threshold selection. The selection of threshold is independent of the solving process of u. Therefore, when changing the threshold, there is no need to recalculate u, which greatly reduces the running time of our algorithm.

To standardize the discussion, we normalize the pixel values g of an image to the [0,1] interval. We achieve this by using the following linear stretching formula:

$$\bar{g} = \frac{g - g_{\min}}{g_{\max} - g_{\min}}.$$

The average value of each region is denoted as $\hat{\rho}_1, \rho_2, \cdots, \rho_K$, and $\hat{\rho}_1 \le \rho_2 \le \cdots \le \rho_K$, is denoted as K-1

$$\rho_i = \frac{\hat{\rho}_i + \rho_{i+1}}{2}, i = 1, 2, \cdots, K-1.$$

Therefore, the i-th($2 \le i \le K-1$)region of an image is $\{x \in \Omega : \rho_{i-1} \le \bar{g}(x) \le \rho_i\}$, the first region is $\{x \in \Omega : 0 \le \bar{g}(x) \le \rho_i\}$, and the Kth region is $\{x \in \Omega : \rho_{i-1} \le \bar{g}(x) \le 1\}$. To obtain the boundary of the i-th($1 \le i \le K$) region, we first set the pixels in the i-th region to 1 and all other pixels to 0; Then call the command contour in Matlab.

4. Numerical experiment

To verify the effectiveness and advantages of our proposed algorithm, we conducted numerical experiments. All numerical experiments were conducted in MATLAB (R2018a) and pyCharm Community Edition 2022.2.1 software on a laptop with a 2.50GHz 12th Gen Intel (R) Core (TM) i5-12500 chip and 16.0GB memory.

We first test the robustness of our method on shape_clean images by adding different types of blur and levels of noise to clean images for image segmentation, and then observe the segmentation effect. Then, we conducted tests on 77200701, DSCN1205_cropped, and sh-med-bandages-082404-01 images in the Weizmann dataset [5] to quantitatively compare the segmentation performance of our algorithm with state-of-the-art image segmentation algorithms (CV [3], SaT [1], T-ROF [4]).

In numerical experiments, we mainly use the quantitative evaluation criterion of F1 score. The formula for calculating F1 is as follows:

$$F1 = \frac{2PR}{P+R},$$

Among them, accuracy rate(Precision) $p = \frac{TP}{TP+FP}$, Recall(Recall) $R = \frac{TP}{TP+FN}$, TP (True Positive) represents the number of pixels correctly predicted as positive by the model, TN (True Negative) represents

ISSN: 2424-8460(O) 2251-2608(P)

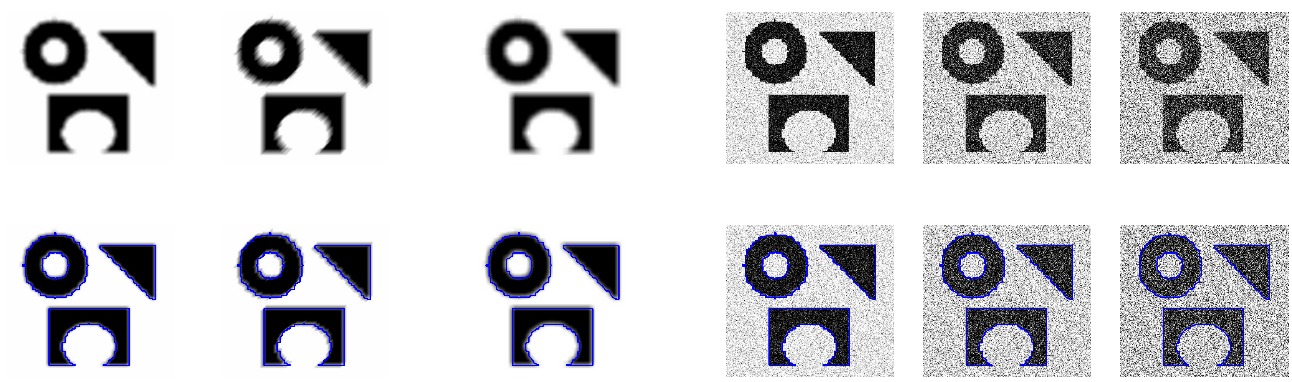
the number of pixels correctly predicted as negative by the model, FP (False Positive) represents the number of pixels incorrectly predicted as positive by the model, and FN (False Negative) represents the number of pixels incorrectly predicted as negative by the model.

Experiment 1: Image segmentation under different types of blur and noise levels

To demonstrate the robustness of our method, we tested it on three blurry and three noisy images, and the test results are shown in **Table 1**. The first row of images are the shape_clean images with added Bd=fspecific (‘Disk ‘, 5), Bm=fspecific (‘motion’, 15, 45), Bg=fspecific (‘gaussian ‘, 12, 12) blurred images and Gaussian noise with variances of 0.3, 0.5, 0.7. The second row of images corresponds to the segmentation results calculated by our algorithm. From the segmentation results, the boundary lines are relatively clear and not affected by blurring or noise. Therefore, our method has stability and robustness under different types of blur and levels of noise.

Table 1. Image segmentation performance under different types of blur and noise levels

Disk Vague	motionVague	gaussianVague	$\sigma^2 = 0.3$	$\sigma^2 = 0.5$	$\sigma^2 = 0.7$



Experiment 2: Comparison of our method with CV, T-ROF, and SaT for real image segmentation performance

In this experiment, the test images are 77200701, DSCN1205_cropped, sh-med-bandages-082404-01 images from the weizmann dataset. First, the same blur and Gaussian noise with a variance of 2.55 are added to the images, and then CV[3], Sat[1], T-ROF[4] and our methods are used to segment the images. The results of image segmentation in this experiment are shown in **Table 2**. From **Table 2**, it can be seen that our method has a higher F1 score than CV [3], SaT [1], and T-ROF [4], which validates the superiority of our algorithm.

Table 2. Comparison of F1 scores between our method and CV, T-ROF, SaT for real image segmentation.

	CV	SaT	T-ROF	Ours
77200701	0.9138	0.9043	0.8532	0.9259
DSCN1205_cropped	0.9555	0.9571	0.9509	0.9611
sh-med-bandages-082404-01	0.9701	0.9715	0.9564	0.9775

5. Conclusion

This article proposes a two-stage strategy image segmentation model and algorithm based on deep plug and play. In the first stage, the energy functional with implicit regularization terms is minimized to obtain a smooth image u. The solution of the w-subproblem uses a deep learning based denoising algorithm. In the second stage, appropriate threshold processing is applied to u to complete segmentation. By conducting numerical experiments on pyCharm and MATLAB, we have verified the robustness and superiority of our proposed algorithm.

References

[1] X. Cai, R. Chan, T. Zeng, A two-stage image segmentation method using a convex variant of the Mumford-Shah model and thresholding, SIAM . Imaging Sci.6 (1)(2013) 368-390.

[2] Zhang K, Li Y, Zuo W, et al. Plug-and-play image restoration with deep denoiser prior[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021.

[3] T.F. Chan, L.A. Vese, Active contours without edges, IEEE Trans. Image Process 10 (2)(2001) 266-277.

[4] X. Cai, R. Chan, C-B. Schonlieb, G. Steidl, T. Zeng, Linkage between piecewise constant Mumford-Shah model and Rudin-Osher-Fatemi model and its virtue in image segmentation, SIAM J. Sci. Comput, 41 (6) (2019) B1310-B1340. 22]

[5] http://www.wisdom.weizmann.ac.il/~vision/Seg_Evaluation_DB/dl.html.