

Research Article

Medical Image Segmentation Algorithm Based on Optimized Convolutional Neural Network-Adaptive Dropout Depth Calculation

Feng-Ping An  ^{1,2} and Jun-e Liu  ³

¹School of Physics and Electronic Electrical Engineering, Huaiyin Normal University, Huaian, JS 223300, China

²School of Information and Electronics, Beijing Institute of Technology, Beijing, BJ 100081, China

³School of Information, Beijing Wuzi University, Beijing, BJ 100081, China

Correspondence should be addressed to Feng-Ping An; anfengping@163.com and Jun-e Liu; 2924175349@qq.com

Received 18 December 2019; Revised 8 March 2020; Accepted 4 May 2020; Published 15 May 2020

Academic Editor: Alex Alexandridis

Copyright © 2020 Feng-Ping An and Jun-e Liu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Medical image segmentation is a key technology for image guidance. Therefore, the advantages and disadvantages of image segmentation play an important role in image-guided surgery. Traditional machine learning methods have achieved certain beneficial effects in medical image segmentation, but they have problems such as low classification accuracy and poor robustness. Deep learning theory has good generalizability and feature extraction ability, which provides a new idea for solving medical image segmentation problems. However, deep learning has problems in terms of its application to medical image segmentation: one is that the deep learning network structure cannot be constructed according to medical image characteristics; the other is that the generalizability γ of the deep learning model is weak. To address these issues, this paper first adapts a neural network to medical image features by adding cross-layer connections to a traditional convolutional neural network. In addition, an optimized convolutional neural network model is established. The optimized convolutional neural network model can segment medical images using the features of two scales simultaneously. At the same time, to solve the generalizability problem of the deep learning model, an adaptive distribution function is designed according to the position of the hidden layer, and then the activation probability of each layer of neurons is set. This enhances the generalizability of the dropout model, and an adaptive dropout model is proposed. This model better addresses the problem of the weak generalizability of deep learning models. Based on the above ideas, this paper proposes a medical image segmentation algorithm based on an optimized convolutional neural network with adaptive dropout depth calculation. An ultrasonic tomographic image and lumbar CT medical image were separately segmented by the method of this paper. The experimental results show that not only are the segmentation effects of the proposed method improved compared with those of the traditional machine learning and other deep learning methods but also the method has a high adaptive segmentation ability for various medical images. The research work in this paper provides a new perspective for research on medical image segmentation.

1. Introduction

At present, the demand for medical imaging in image-guided radiotherapy, image-guided surgery, image-guided interventional therapy, and image-guided navigation is increasing, which promotes the research and development of medical imaging technology and image processing technology [1, 2]. The accuracy and efficiency of image-guided surgery is much higher than that of traditional surgical

procedures, and it can reduce the risk of surgery. Medical image segmentation is the key technology for image guidance. The advantages and disadvantages of image segmentation play an important role in image-guided surgery [3–5]. Medical image segmentation is a complex and critical step in the field of medical image processing and analysis. Its purpose is to depict the anatomical structure of interest or the area of a particular tissue as accurately as possible. Although there are many kinds of medical image

segmentation methods, medical image segmentation methods are mainly divided into the following two categories: medical image segmentation methods based on traditional machine learning and medical image segmentation methods based on deep learning [6–9].

The medical image segmentation methods based on traditional machine learning mainly include the following: Rodrigues et al. [10] proposed to use the support vector machine method to detect the region of interest and then used the Adaboosting weak classifier to select the features on the region of interest. Finally, the segmentation task is completed based on the active contour model, which achieves certain effects. However, the feature extraction effect of the machine learning method is poor, which affects the image segmentation effect in the later stage. Anter et al. [11, 12] proposed a medical image segmentation method based on fuzzy mean, which achieved good results in medical image segmentation experiments. However, such methods have weak adaptive capabilities. Huang et al. [13] proposed a graph-based breast ultrasound segmentation method that combines regional statistical information. It has achieved certain effects in breast ultrasound image segmentation, but its stability is poor. Yu et al. [14] proposed different forms of 3DGrabCut to perform three-dimensional breast ultrasound segmentation tasks and achieved good results in the correlation of breast ultrasound image segmentation, but the adaptive ability was weak. Guo et al. [15] proposed a deformation method based on dictionary learning, which improved the learning strategy on the existing dictionary learning model and then performed image segmentation. It has achieved certain effects in medical image segmentation, but the classification accuracy is low.

Due to the above problems for medical image segmentation methods based on traditional machine learning, many scholars continue to explore new medical image segmentation methods. In recent years, deep learning [16, 17] has received extensive attention in many applications of artificial intelligence. It has achieved a quantum leap in precision in a variety of applications. It provides a feasible way to solve the above problems. It is in this technical background that a second type of medical image segmentation method based on deep learning was proposed. Ronneberger et al. [18] proposed a computational segmentation model based on a full convolutional network. It reconstructs the reduced-dimensional image by deconvolution and combines the information of the corresponding layer of the convolutional dimension reduction to the inverse by means of hopping. Each convolution layer corresponds to a deconvolution layer. Therefore, automatic segmentation of medical images is achieved. Milletari et al. [19] calculated the improvement of U-Net and studied the use of three-dimensional convolution. It has achieved good segmentation results in medical image segmentation. Xu et al. [20] proposed a multitissue target segmentation of breast tumors based on convolutional neural networks, which includes skin, fibrous tissue, tumor, and adipose tissue. Kumar et al. [21] proposed a segmentation and detection task for breast tumor mass based on a full

convolutional network model. Ahmad et al. [22] proposed the use of deep stacking automatic encoder for liver segmentation of CT images. The classification accuracy of liver segmentation is 91%. Tang et al. [23] proposed using a fast convolutional neural network to segment the liver image. First, the rapid convolutional neural network was used to locate the liver region, and then the detection results were input into the depth frame to segment the liver image. Hu et al. [24] proposed a new automatic liver segmentation method based on a deep three-dimensional convolutional neural network and global optimization of the surface evolution. Yang et al. [25] used 2D generation to constrain the network to segment the liver image. To further improve segmentation accuracy, Li et al. [26] proposed the use of the 3D method to segment the liver and liver tumors, combined with the 3D fully convolutional network and graph cut algorithm to achieve automatic segmentation in CT images. Sun et al. [27] used multiple criteria to enhance different features of CT images and proposed multicore fully convolutional network (MK-FCN) automatic liver tumor segmentation. This technique uses multiple ratios to enhance different features of CT images for multiphase high-order feature fusion. In [28], a method of combining traditional deformation models and convolutional networks was proposed to achieve accurate segmentation of the prostate. To make full use of the global spatial information of the image and overcome the large computational complexity of the patch-based segmentation algorithm, Karimi et al. [29] used fully convolutional network training to achieve prostate segmentation. Milletari et al. [19] proposed a new fully convolutional network based on the traditional fully convolutional network. This network model introduces the residual network idea in the feature extraction layer of the U-Net network and solves the sample imbalance problem with the similarity coefficient. Yu et al. [30] proposed a convolutional model based on residual thought. The long and short residual connection method not only integrates local and global feature information well but also improves the model training efficiency and recognition ability. However, it has a more serious overfitting problem. For the above reasons, the deep learning model has been deeply applied and popularized in medical image segmentation. However, the deep learning method has the following problems in the application process. (1) Deep learning model network structure optimization problem: in a standard convolutional neural network, each hidden layer can only obtain input from its neighboring previous layer and then pass its output to the next lower layer. This not only limits the flexibility of the convolutional neural network structure but also does not comprehensively utilize multiscale feature information. (2) Overfitting problem: the overfitting phenomenon severely reduces the effectiveness of the deep learning model in learning heterogeneous data features, which reduces the classification accuracy of the heterogeneous data by the depth calculation model. At the same time, the overfitting phenomenon limits the number of hidden layers in the deep learning model. The traditional dropout method can reduce the occurrence of overfitting. However, it limits the generalization ability of deep learning models

[31–35]. In view of this, to better solve the network structure problem of deep learning model, this paper establishes an optimized convolutional neural network model by adding cross-layer connections in traditional convolutional neural networks. In the optimized convolutional neural network model, features from different scales can be used simultaneously for segmentation. At the same time, in order to solve the problem that the traditional dropout method reduces the generalizability of the deep learning model, this paper proposes an adaptive dropout model. According to the position of the hidden layer, an adaptive distribution function is designed to set the probability that the neurons of each layer are activated. It further enhances the generalization capabilities of the dropout model. Then, the adaptive distribution function is applied to the deep learning model, and a deep learning model based on adaptive dropout is proposed. Based on the above ideas, this paper proposes a medical image segmentation algorithm based on optimized convolutional neural network-adaptive dropout depth calculation.

Section 2 of this paper focuses on the optimization of the convolutional neural network framework model. Section 3 systematically describes the depth calculation model based on adaptive dropout proposed in this paper. Section 4 introduces a medical image segmentation algorithm based on optimized convolutional neural network-adaptive dropout depth calculation. Section 5 analyzes the medical image segmentation algorithm proposed in this paper and compares it with the mainstream medical image segmentation algorithm. Finally, the paper is summarized and discussed.

2. Optimized Convolutional Neural Network Model

2.1. Optimized Convolutional Neural Network Model Framework. The optimized convolutional neural network model framework proposed in this section includes an input layer, an s -group interleaved convolutional layer and pooling layer, a fully connected layer, and an output layer. The fully connected layer is obtained by splicing these convolutional layers and pooling layers with a cross-layer connection, as shown specifically in Figure 1.

The input x is a 3-dimensional array of size $h \times w \times n$, where h and w are spatial dimensions, n is the channel dimension, $n=3$ represents a color image, and $n=1$ represents a grayscale image. We use “ $*$ ” to indicate the convolution operation and “ f ” to indicate the activation function, and the convolutional layer is calculated as follows:

$$h_{2k-1,j}^l = c_{k,j}^l = f(u_{2k-1,j}^l) = f\left(\sum_i h_{2k-2,i}^l * W_{ij}^{2k-1} + b_j^{2k-1}\right), \quad 1 \leq k \leq s, \quad (1)$$

where W_{ij}^{2k-1} represents a weight matrix between the i th feature face of the $(2k-2)$ th hidden layer and the j th feature face of the $(2k-1)$ th hidden layer. b_j^{2k-1} represents the offset of the j th feature plane of the $(2k-1)$ th hidden layer. $c_{k,j}^l$ represents the j th feature plane of the k th convolutional layer. $h_{2k-1,j}^l$ and $h_{2k-2,i}^l$ represent the i th feature face of the

$(2k-2)$ th hidden layer of the 1 st sample and the j th feature face of the $(2k-1)$ th hidden layer, respectively. The activation function f can be a sigmoid or a modified linear unit. Here, let $h_0^l = x^l$.

In each pooled layer, a fixed step size is used for all feature faces. The pooling function can be expressed as

$$h_{2k,j}^l = t_{k,j}^l = \text{pooling}\{h_{2k-1,j}^l\}, \quad 1 \leq k \leq s. \quad (2)$$

In the formula, $\text{pooling}\{\cdot\}$ can be average pooling or maximum pooling. $h_{2k,j}^l$ and $h_{2k-1,j}^l$ represent the j th feature face of the $(2k-1)$ th hidden layer of the l th sample and the j th feature face of the 2 th hidden layer, respectively. $t_{k,j}^l$ represents the j th feature plane of the k th pooling layer of the l th sample.

A fully connected layer is a splicing obtained by the activation of two or more convolutional layers and a pooled layer through a cross-layer connection. It can form the entire multiscale feature discriminant vector. In fact, the fully connected layer has the following form:

$$h_{2s+1}^l = (a_1 h_1^l, a_2 h_2^l, \dots, a_{2s} h_{2s}^l), \quad (3)$$

where h_{2k-1}^l and h_{2k}^l represent the $(2k-1)$ th hidden layer and the $2k$ th hidden layer of the l th sample, respectively. Let the binary string $SI = a_1, a_2, \dots, a_{2s-1}$ denote a crossover indicator that indicates how the cross-layer connection is made. For example, the cross-layer connection method indicated by $SI = 111\dots 1$ is that all $2s-1$ convolution layers and pooling layers are connected to the fully connected layer. The cross-layer connection method indicated by $SI = 100\dots 0$ is that only the first convolutional layer is connected to the fully connected layer. The cross-layer connection method indicated by $SI = 000\dots 0$ is that there is no cross-layer connection, and this is the standard CNN. The actual output is a C -dimensional softmax, which predicts the probability distribution of C different categories, expressed as

$$o^l = \text{soft max}(u^l) = \text{soft max}(W^{2s+2} h_{2s+1}^l + b^{2s+2}), \quad (4)$$

where W^{2s+2} and b^{2s+2} represent the weight and offset of the output layer, respectively, and $\text{soft max}_i(x) = \exp(x_i)/\sum_j \exp(x_j)$.

2.2. Learning Algorithm for Optimizing Convolutional Neural Network Framework. For the l th sample, the optimized convolutional neural network algorithm proposed in this paper uses the following formula to sequentially calculate the activation of all convolutional and pooled, fully connected, and output layers:

$$\begin{cases} h_{2k-1,j}^l = f(u_{2k-1,j}^l) = f\left(\sum_i h_{2k-2,i}^l * W_{ij}^{2k-1} + b_j^{2k-1}\right), & 1 \leq k \leq s, \\ h_{2k,j}^l = \text{pooling}\{h_{2k-1,j}^l\}, & 1 \leq k \leq s, \\ h_{2s+1}^l = (a_1 h_1^l, a_2 h_2^l, \dots, a_{2s} h_{2s}^l), \\ o^l = \text{soft max}(u^l) = \text{soft max}(W^{2s+2} h_{2s+1}^l + b^{2s+2}). \end{cases} \quad (5)$$

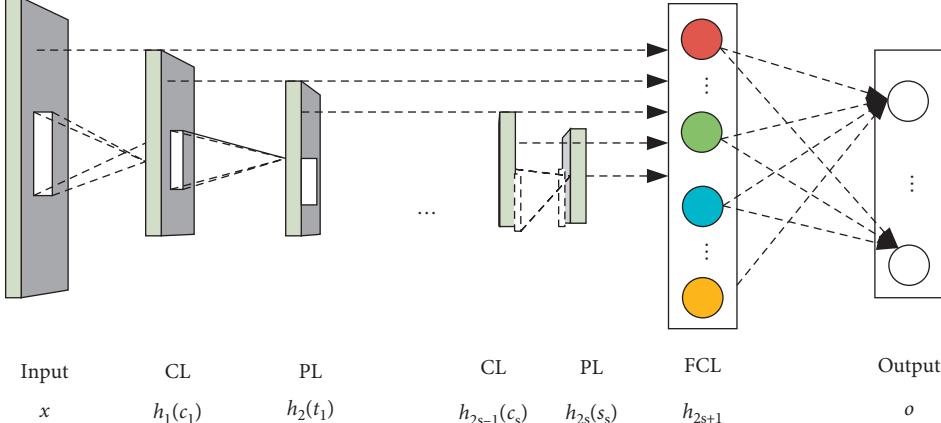


FIGURE 1: Splicing a convolutional neural network. In the framework, the features of different convolutional layers and pooled layers are spliced to obtain a fully connected layer, and the fully connected layer is directly input to the output layer.

Let $y^l = (y_1^l, y_2^l, \dots, y_C^l)^T$ denote the desired output. $o^l = (o_1^l, o_2^l, \dots, o_C^l)^T$ represents the actual output. We use cross entropy loss as the objective function, specifically:

$$L_N(y^l, o^l) = -\sum_{l=1}^N \sum_{c=1}^C y_c^l \log(o_c^l). \quad (6)$$

First, we calculate the sensitivity δ_k^l , $1 \leq k \leq 2s + 1$ of each hidden layer and the sensitivity δ^l of the output layer, specifically:

$$\left\{ \begin{array}{l} \delta^l = o^l - y^l, \\ \delta_{2s+1}^l = [(W^{2s+2})^T \delta^l] \circ \text{soft max}'(u_{2s+2}^l), \\ (\delta_{1,FC}^l, \delta_{2,FC}^l, \dots, \delta_{2k,FC}^l, \dots, \delta_{2s,FC}^l) = \delta_{2s+1}^l, \quad 1 \leq k \leq s, \\ \delta_{2s}^l = \delta_{2s,FC}^l, \\ \delta_{2k-1,j}^l = f'(u_{k,j}^l) \circ \text{uppooling}\{\delta_{2k,j}^l\} + a_{2k-1} \delta_{2k-1,FC,j}^l, \\ \delta_{2k,j}^l = \delta_{2k+1,j}^l * \text{rot180}(W^{2k+1}) + a_{2k} \delta_{2k,FC,j}^l, \end{array} \right. \quad (7)$$

where δ^l represents the sensitivity of the output layer. δ_{2k-1}^l and δ_{2k}^l represent the sensitivity of the $(2k-1)$ th hidden layer and the $2k$ th hidden layer, respectively. $\delta_{2k-1,FC}^l$ and $\delta_{2k,FC}^l$ are part of the $(2k+1)$ th hidden layer, that is, the fully connected layer, corresponding to the $(2k-1)$ th and the $2k$ th hidden layer, respectively. In addition, $\text{uppooling}\{\bullet\}$ is an upsampling function of the pooling function defined by equation (2). $\text{Softmax}'(\bullet)$ represents the derivative of the soft maximum function. $\text{rot180}(\bullet)$ indicates horizontal and vertical flipping operations on the matrix. The symbol “ $*$ ” indicates the Hadamard product.

The weights of the optimized convolutional neural network model and the inverse of the offset can be calculated by formulas (6) and (7), as follows:

$$\left\{ \begin{array}{l} \frac{\partial L_N}{\partial W^{2s+2}} = \sum_{l=1}^N \delta^l (h_{2s+1}^l)^T, \\ \frac{\partial L_N}{\partial b^{2s+2}} = \sum_{l=1}^N \delta^l, \\ \frac{\partial L_N}{\partial W_{ij}^{2k-1}} = \sum_{l=1}^N \delta_{2k-1,j}^l * h_{2k-2,i}^l, \\ \frac{\partial L_N}{\partial b_j^{2k-1}} = \sum_{l=1}^N \delta_{2k-1,j}^l, \\ 1 \leq k \leq s. \end{array} \right. \quad (8)$$

Based on formulas (5) to (8), this paper proposes a gradient descent learning algorithm that conforms to the optimized convolutional neural network model proposed in this section. That is, the cross-layer backpropagation algorithm of the convolutional neural network model is optimized.

This section proposes a stitching framework for cross-continuous convolutional neural networks. The framework consists of an input layer, a convolutional layer, a pooling layer, a fully connected layer, and an output layer. In this framework, the fully connected layer is obtained by concatenating one or more convolutional layers and pooling layers in front of it using a fixed weight cross-layer connection. The framework uses the features of the nonhighest convolutional layer or pooling layer and then stitches to obtain multiscale features for segmentation or recognition. It has good stability to different pooling mechanisms, initialization methods, optimization methods, activation functions, and convolution kernels of different numbers and sizes.

2.3. Description of Image Analysis with Full Connection

- (1) Conversion idea of fully connected layer and convolutional layer

The difference between the fully connected layer and the convolutional layer is that the neuron in the convolutional layer is only connected to a local region in the input data, and the neurons in the convolutional column share parameters. But in these two types of layers, neurons calculate dot products. So, their functional form is the same. Therefore, the two can be transformed into each other.

For any convolutional layer, it has a fully connected layer with the same forward propagation function. The weight matrix is a huge matrix, except for some specific blocks, and the rest are zero. The elements in most of the blocks are equal. Similarly, any fully connected layer can be converted into a convolutional layer. For example, for a fully connected layer with $K=2048$, the input size is $7 \times 7 \times 256$. The fully connected layer can be regarded as a convolutional layer with $F=7$, $P=0$, $S=1$, $K=2048$. So, it can directly analyze or manipulate the image.

- (2) An example of the conversion of a fully connected layer to a convolutional layer

Suppose that an input image of $224 \times 224 \times 3$ of a convolutional neural network, a series of convolutional layers, and downsampling layers transform the image data into an active data volume of size $7 \times 7 \times 512$. Here, two fully connected layers of size 2048 can be used, and the last fully connected layer with 1000 neurons is used to calculate the image segmentation. It can convert any of these three fully connected layers into a convolutional layer. The specific steps are as follows:

- (1) For the fully connected layer whose first connection area is $(7 \times 7 \times 256)$, let its filter size be $F=7$. In this way, the output data volume is $(1 \times 1 \times 2048)$.
- (2) For the second fully connected layer, let $F=1$. In this way, the output data volume is $(1 \times 1 \times 2048)$.
- (3) Perform the similar operation for the last fully connected layer; let $F=1$. The final output is $(1 \times 1 \times 1000)$.

3. Depth Calculation Model Based on Adaptive Dropout

Two key problems in the design of adaptive dropout depth calculation model are adaptive distribution function design and adaptive high-order backpropagation algorithm design.

3.1. Adaptive Distribution Function Design. In this paper, an adaptive distribution function is designed to hide the position of the layer as an independent variable and to hide the activation probability of the layer neurons as the dependent variable. The probability of neuron activation for each hidden layer is set according to the position of the hidden layer. Through the above analysis, the adaptive distribution function monotonically decreases with the increase of the hidden layer position [31, 32]. In addition, the dropout model proposed by Hinton shows that the activation probability of neurons in each layer should be around 0.5. This ensures that enough activation neurons are retained. At the same time, enough neurons are discarded to improve the generalization ability of the depth calculation model. Finally, the value of the dropout rate for each layer needs to be in the interval $(0, 1)$. Therefore, the adaptive distribution function designed in this paper is as follows:

$$\rho = f(l) = \begin{cases} 1 - \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^l \exp\left(-\frac{(1-n/2)^2}{2\sigma^2}\right) dl, & n = 2k (k \in N_+), \\ 1 - \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^l \exp\left(-\frac{(l-(n+1)/2)^2}{2\sigma^2}\right) dl, & n = 2k + 1 (k \in N_+), \end{cases} \quad (9)$$

where the dropout rate is given, l represents the position of the hidden layer, n represents the number of layers of the depth calculation model, and σ represents the parameter used to control the range of the dropout rate. The adaptive function $\rho=f(l)$ has the following properties:

- (1) When the hidden layer position increases, the probability of the activation of neurons is reduced. The proof is as follows.

The derivative of the function $f(l)$ versus l is as follows:

$$f'(l) = \frac{d\rho}{dl} \begin{cases} -\frac{1}{\sigma\sqrt{2\pi}} e^{-(ln/2)^2/2\sigma^2} < 0, & n = 2k (k \in N_+), \\ -\frac{1}{\sigma\sqrt{2\pi}} e^{-(l(n+1)/2)^2/2\sigma^2} < 0, & n = 2k + 1 (k \in N_+). \end{cases} \quad (10)$$

Since $f'(l) < 0$, $\rho=f(l)$ is a monotonically decreasing function.

(2) The probability of the activation of neurons in the middle hidden layer is 0.5.

The proof is as follows.

When the number of hidden layers is even, that is, $n = 2k (k \in N_+)$, we obtain

$$\rho = f(n/2) = \left\{ 1 - \int_{-\infty}^{n/2} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(l-n/2)^2}{2\sigma^2}\right) dl \right\} = 0.5. \quad (11)$$

When the number of hidden layers is odd, that is, $n = 2k + 1 (k \in N_+)$, we obtain

$$\rho = f((n+1)/2) = \left\{ 1 - \int_{-\infty}^{(n+1)/2} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(l-(n+1)/2)^2}{2\sigma^2}\right) dl \right\} = 0.5. \quad (12)$$

Therefore, the activation probability of each neuron in an intermediate hidden layer is 0.5.

(3) $\rho \in (0, 1)$, that is, the probability that each neuron is activated is between (0, 1).

The proof is as follows.

$f(l)$ is a strictly monotonically decreasing function, so we have

$$f(n) < f(l) < f(1). \quad (13)$$

When the number of hidden layers is even, that is, $n = 2k (k \in N_+)$, we obtain

$$\begin{aligned} \rho_{\max} &= f(l) = 1 - \int_{-\infty}^l \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(l-n/2)^2}{2\sigma^2}\right) dl \\ &= 1 - \left(\lim_{x \rightarrow -\infty} \left(\frac{1}{2} \operatorname{erf}\left(\frac{\sqrt{2}(-2l+n)}{4\sigma}\right) - \frac{1}{2} \operatorname{erf}\left(\frac{\sqrt{2}(-2+n)}{4\sigma}\right) \right) \right) < 1, \end{aligned} \quad (14)$$

$$\begin{aligned} \rho_{\min} &= f(n) = 1 - \int_{-\infty}^n \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(l-n/2)^2}{2\sigma^2}\right) dl \\ &= 1 - \left(\lim_{x \rightarrow -\infty} \left(\frac{1}{2} \operatorname{erf}\left(\frac{\sqrt{2}(-2l+n)}{4\sigma}\right) + \frac{1}{2} \operatorname{erf}\left(\frac{n\sqrt{2}}{4\sigma}\right) \right) \right) > 0. \end{aligned} \quad (15)$$

When the number of hidden layers is odd, that is, $n = 2k + 1 (k \in N_+)$, we obtain

$$\begin{aligned} \rho_{\max} &= f(l) = 1 - \int_{-\infty}^l \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(l-(n+1)/2)^2}{2\sigma^2}\right) dl \\ &= 1 - \left(\lim_{x \rightarrow -\infty} \left(\frac{1}{2} \operatorname{erf}\left(\frac{\sqrt{2}(-2l+n+1)}{4\sigma}\right) - \frac{1}{2} \operatorname{erf}\left(\frac{\sqrt{2}(-1+n)}{4\sigma}\right) \right) \right) < 1, \end{aligned} \quad (16)$$

$$\begin{aligned} \rho_{\min} &= f(n) = 1 - \int_{-\infty}^n \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(l-(n+1)/2)^2}{2\sigma^2}\right) dl \\ &= 1 - \left(\lim_{x \rightarrow -\infty} \left(\frac{1}{2} \operatorname{erf}\left(\frac{\sqrt{2}(-2l+n+1)}{4\sigma}\right) + \frac{1}{2} \operatorname{erf}\left(\frac{(n-1)\sqrt{2}}{4\sigma}\right) \right) \right) > 0. \end{aligned} \quad (17)$$

From formulas (14) to (17), it can be seen that for any l , $f(l) \in (0, 1)$ has a probability of each neuron being activated that is in (0, 1).

3.2. Depth Calculation Model Based on Adaptive Dropout.
Let L denote the total number of layers based on the adaptive dropout depth calculation model. The number of hidden

layers is $L - 2$, and $l \in \{0, 1, \dots, L - 1\}$ is the index number of each layer, that is, $l = 0$. For the input layer, $l = 1$ is the first hidden layer, so $l = L - 1$ is the output layer. For the depth calculation model represented by the N th order tensor of each layer, it is assumed that $z_{j_1 j_2 \dots j_N}^{(l)}$ is the input of the first-layer neurons j_1, j_2, \dots, j_N , and $y_{j_1 j_2 \dots j_N}^{(l)}$ is the output of the first-layer neurons j_1, j_2, \dots, j_N . $y^{(0)} = x$ is the input data and the output of the $y^{(L-1)}$ model. $W^{(l)}$ and $b^{(l)}$ are the weight tensor and offset tensor of the l th layer, respectively. Therefore, the forward propagation process of this depth calculation model is as follows:

$$\begin{aligned} z_{j_1 j_2 \dots j_N}^{(l+1)} &= W_\alpha^{(l)} \odot Y^{(l)} + b_{j_1 j_2 \dots j_N}^{(l)}, \\ \alpha &= j_n + \sum_{i=1}^{N-1} (j_i - 1) \prod_{t=i+1}^N J_t, \end{aligned} \quad (18)$$

$$y_{j_1 j_2 \dots j_N}^{(l+1)} = f(z_{j_1 j_2 \dots j_N}^{(l)}), \quad (19)$$

where \odot denotes the multipoint product of two tensors; let $H \in R^{J_1 \times J_2 \times \dots \times J_N}, J_1 \times J_2 \times \dots \times J_N = \alpha$ denote the result of the multipoint product of $W \in R^{\alpha \times J_1 \times J_2 \times \dots \times J_N}$ and tensor $A \in R^{J_1 \times J_2 \times \dots \times J_N}$. Each element in H is defined as follows:

$$h_{j_1 j_2 \dots j_N} = W_\beta \cdot A, \beta = j_n + \sum (j_t - 1) \prod_{t=i+1}^N J_t, \quad (20)$$

where f is the activation function. The depth calculation model based on adaptive dropout proposed in this paper uses the sigmoid function as the activation function of neurons; that is, $f(x) = 1/(1 + \exp(-x))$.

The forward propagation process for the depth calculation model based on adaptive dropout is as follows:

$$r_{j_1 j_2 \dots j_N}^{(l)} \sim \text{Bernoulli}(\rho^{(l)}), \quad (21)$$

$$\tilde{y}_{j_1 j_2 \dots j_N}^{(l)} = r_{j_1 j_2 \dots j_N}^{(l)} \cdot y_{j_1 j_2 \dots j_N}^{(l)}, \quad (22)$$

$$\begin{aligned} z_{j_1 j_2 \dots j_N}^{(l+1)} &= W_\alpha^{(l)} \odot \tilde{Y}^{(l)} + b_{j_1 j_2 \dots j_N}^{(l)}, \\ \alpha &= j_n + \sum_{i=1}^{N-1} (j_i - 1) \prod_{t=i+1}^N J_t, \end{aligned} \quad (23)$$

$$y_{j_1 j_2 \dots j_N}^{(l+1)} = f(z_{j_1 j_2 \dots j_N}^{(l)}), \quad (24)$$

where $\rho^{(l)}$ is the activation probability of each hidden layer neuron in the l th layer. $r_{j_1 j_2 \dots j_N}^{(l)}$ obeys a Bernoulli distribution, and its probability $\rho^{(l)}$ takes a value of 1. For the depth calculation model based on adaptive dropout, the probability of activation of each hidden layer neuron is set by an adaptive function, and a submodel of a depth calculation model can be obtained for each training. The adaptive dropout model enhances the generalizability of the deep computation model by training multiple submodels with shared weights.

3.3. Adaptive High-Order Backpropagation Learning Algorithm.

The reconstruction error function based on the

adaptive dropout depth calculation model is defined as follows:

$$J_{TAE}(\theta; l) = 0.5 \left(h_{W,b}(\tilde{y}^{(l-1)}) - \tilde{y}^{(l-1)} \right)^T G \left(h_{W,b}(\tilde{y}^{(l-1)}) - \tilde{y}^{(l-1)} \right), \quad (25)$$

where $\theta = \{W^{(l-1)}, b^{(l-1)}; W^{(l-1)'}, b^{(l-1)'}\}$ is a parameter and $\tilde{y}^{(l-1)}$ is an expansion vector corresponding to the output tensor of the activated neurons of the $l-1$ th layer.

As with the basic depth calculation model, the depth-difference calculation model based on the adaptive dropout uses the gradient descent method to update the parameters of the model. However, for each step of the training process based on the adaptive dropout depth model, forward propagation and backpropagation are trained on a submodel set by the adaptive distribution function. Therefore, when calculating the parameter gradient, only the activated neurons are calculated.

4. Medical Image Segmentation Algorithm Based on Optimized Convolutional Neural Network-Adaptive Dropout Depth Calculation and Example Analysis

4.1. Medical Image Segmentation Algorithm Based on Optimized Convolutional Neural Network-Adaptive Dropout Depth Calculation. Based on the analysis in Sections 2 and 3 of this paper, this section constructs a medical image segmentation algorithm based on an optimized convolutional neural network with adaptive dropout depth calculation. First, this paper establishes an optimized convolutional neural network model by adding cross-layer connections in a traditional convolutional neural network. This can effectively solve the problem of network structure flexibility in a deep learning model. At the same time, this paper proposes an adaptive dropout model, which can effectively enhance the generalizability of the dropout method to reduce the deep learning model. Based on the above analysis, this paper proposes a medical image segmentation algorithm based on an optimized convolutional neural network with adaptive dropout depth calculation. The basic idea of the proposed medical image segmentation algorithm is shown in Figure 2. The basic steps corresponding to this idea are as follows:

- (1) First, the medical image data to be segmented are subjected to preprocessing such as denoising, adding, and expanding.
- (2) To better solve the problem of network structure of deep learning model, this paper establishes an optimized convolutional neural network model by adding cross-layer connections in traditional convolutional neural networks. The optimized convolutional neural network model can simultaneously use two scale features for segmentation, which can better position medical image segmentation services.
- (3) This paper proposes an adaptive dropout model. According to the position of the hidden layer, an adaptive distribution function is designed to set the activation probability of each layer of neurons, which

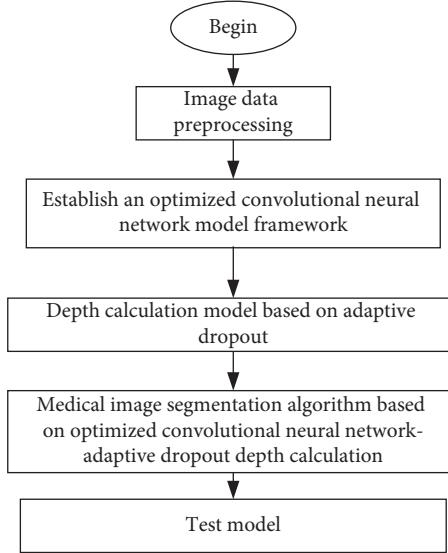


FIGURE 2: Basic idea of medical image segmentation algorithm based on optimized convolutional neural network-adaptive dropout depth calculation.

further improves the generalizability of the dropout model. This solves the problem of the general dropout method reducing the generalizability of the deep learning model. It can further enhance the generalizability of the deep learning model and reduce the risk of overfitting.

- (4) The method of step (3) is introduced into step (2), and a medical image segmentation algorithm is established based on an optimized convolutional neural network with adaptive dropout depth calculation through steps (1)–(3). The algorithm is used to analyze related examples and compare and analyze with other mainstream medical image segmentation methods.

4.2. Ultrasonic Tomography Dataset Experiment. To verify the segmentation effect of the proposed algorithm on medical images, this section will describe segmentation tests on a dataset composed of ultrasonic tomographic images from Delphinus Medical Technologies, USA [36, 37], and compare the proposed algorithm with mainstream medical image segmentation algorithms.

4.2.1. Dataset and Segmentation Process Description. The dataset images for this experiment were derived from ultrasound tomography (UST) volume images from Delphinus Medical Technologies, USA. The resolution of all images is $0.5 \times 0.5 \times 2.0 \text{ mm}^3$, and the matrix size of the coronal plane is 512×512 . All UST images meet the requirements of clinical diagnosis. In this experiment, 50 UST volume data were used, and each volume data contains 36 slice images. Some examples are shown in Figure 3.

The deep learning model used in this experiment was based on the Pytorch implementation and was trained on a

Titan-X GPU. The network architecture in the deep learning model is the network architecture proposed in Section 2 of this paper. The overfitting processing method of the deep learning model is the adaptive depth-based deep computing model proposed in Section 3. The initial learning rate was set to 0.01, and the learning rate was reduced to one-tenth of the original rate when training progressed to 80 and 120 epochs. The training lasted for 200 epochs. In all training sessions, the model proposed in this paper is trained based on the stochastic gradient descent method, and the number of samples per batch is set to 32.

In this experiment, the Dice coefficient (D), Jaccard coefficient (J), and number of false positives (FPs) were used to evaluate the effect of breast segmentation. The specific evaluation indicators are described in [38].

4.2.2. Segmentation Results and Analysis. The image segmentation algorithm proposed in this paper is used to segment the experimental dataset image separately from other mainstream medical image segmentation algorithms. The partial image segmentation results are shown in Figure 4.

Figure 4 shows the segmentation effect of the partial segmentation algorithm, and Figures 4(a)–4(e) represent the reference image, the segmentation results of the methods in [39–41], and the segmentation result of the proposed method, respectively. As seen from Figure 4, methods [39, 40] have segmentation errors, which are marked with blue and red circles, respectively. Visually, there is no significant difference between method [41] and the method of this paper. For better comparison and analysis, Table 1 gives the index comparison data for the specific segmentation results in [39–43].

It can be seen from Table 1 that the segmentation effect of the medical image segmentation algorithm based on the optimized convolutional neural network-adaptive dropout depth calculation is better than that of the traditional machine learning segmentation algorithms proposed in [39, 40] and has a large increase over the segmentation effect of the deep learning algorithms proposed in [41–44]. This illustrates the unique advantages of the algorithm proposed in this paper. Specifically, the traditional machine learning methods proposed in [39, 40] have the lowest Dice values after image segmentation, 0.8397 and 0.8485, respectively, and the Jaccard indexes are lower than 0.8, while the FP indexes are higher than 0.1. This shows that the machine learning method is one of the worst-case methods among the above-listed methods. The image segmentation method proposed in [41] without optimized convolutional neural network model and improved dropout deep learning model has Dice index higher than 0.9, Jaccard index higher than 0.85, and FP index lower than 0.06. It shows that the convolutional neural network model that has not been optimized has a significant improvement over traditional machine learning methods. This is mainly because the deep learning model can better train the experimental data and obtain a more reasonable and reliable image segmentation model. It shows that the deep learning method is suitable for

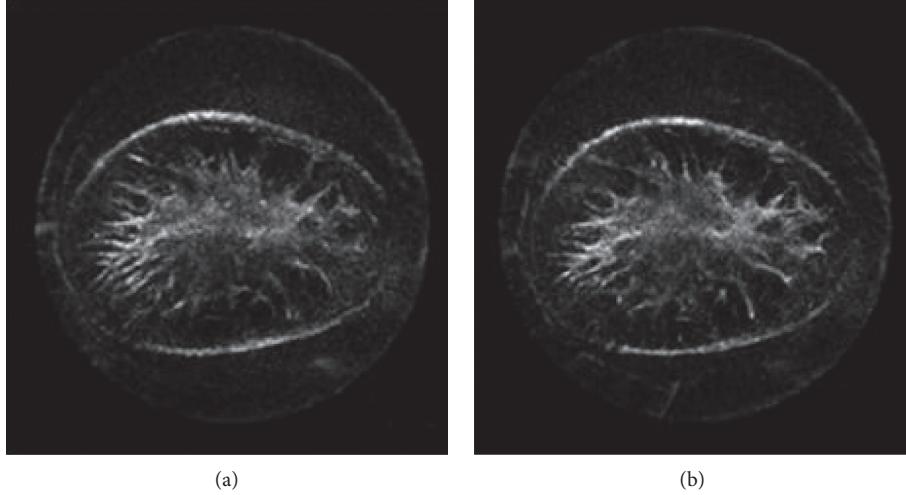


FIGURE 3: UST image part example.

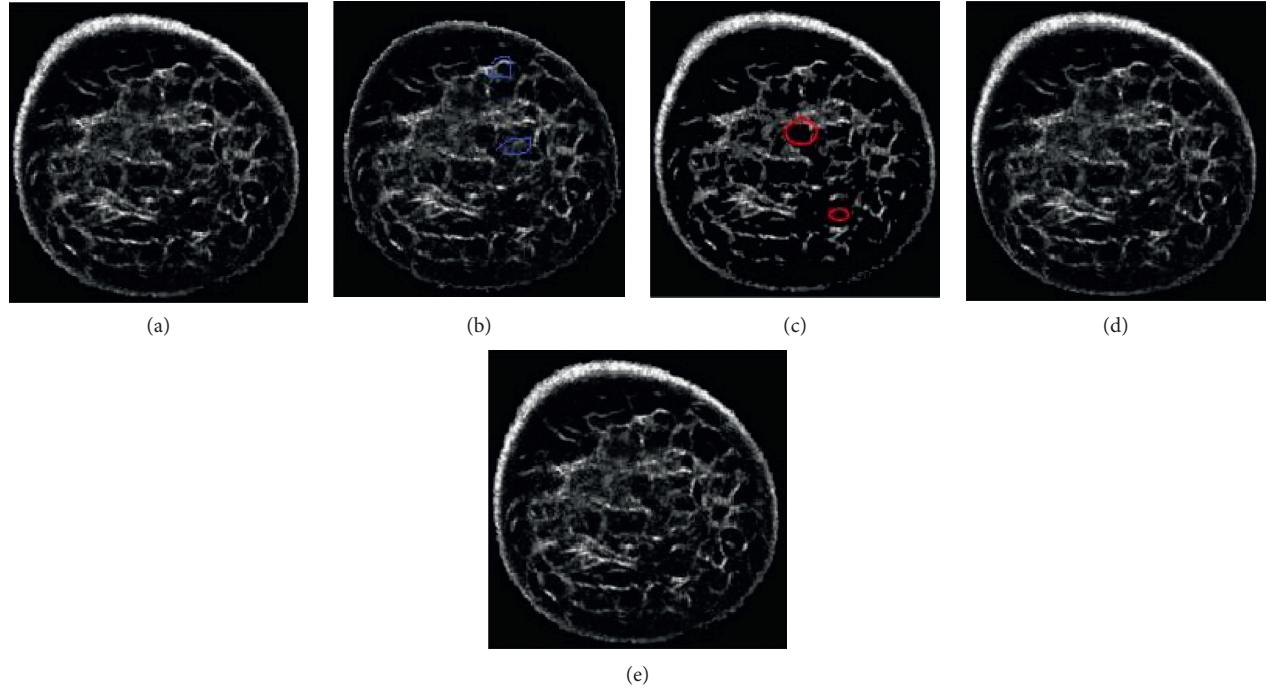


FIGURE 4: Partial image segmentation results. (a) Datum image. (b) [39]. (c) [40]. (d) [41]. (e) Ours.

TABLE 1: Comparison of ultrasonic tomographic image dataset segmentation results.

Segmentation method	Dice	Jaccard	FP
[39]	0.8397	0.7826	0.1337
[40]	0.8485	0.7932	0.1191
[41]	0.9132	0.8557	0.0582
[42]	0.9307	0.9018	0.0311
[43]	0.9402	0.9314	0.0210
[44]	0.9516	0.9432	0.0183
Ours	0.9904	0.9827	0.0012

image segmentation. The deep learning methods proposed in [42–44] have optimized the convolutional neural network model to varying degrees. The Dice and Jaccard indicators obtained by using them for image segmentation are both above 0.9, and the FP indicators are below 0.05. It shows that the methods proposed in [42–44] have not improved the image segmentation effect obviously compared with the traditional machine learning method. Moreover, it has been improved to a certain extent compared to the nonoptimized convolutional neural network model. It shows that the optimization of deep learning models and the improvement of dropout have a greater effect on improving the effect of image segmentation. The Dice and Jaccard indicators for image segmentation obtained by the method proposed in this paper are the highest among those of all the methods, and the FP index is the lowest among those of all the methods. This shows that the proposed method is the best among all the methods. This is mainly because the method proposed in this paper is better optimized than the deep learning methods proposed in [41–43]. The proposed method not only optimizes the network structure of the deep learning model but also solves the overfitting problem of the deep learning model. Through this processing, the deep learning model obtains a stronger modeling ability and nonlinear fitting ability and fully exploits the advantages of the deep learning method in medical image segmentation.

4.3. SpineWeb Dataset Experiment. To further verify the segmentation effect of the proposed algorithm on lumbar CT medical images, this section discusses segmentation tests on the two SpineWeb datasets [45] and compares the results with those of mainstream medical image segmentation algorithms.

4.3.1. The Dataset and Segmentation Process. The dataset of images for this experiment was derived from two datasets from SpineWeb, one of which contained six volumetric images that labeled segmented vertebral bodies without transverse processes, spinous processes, or pedicles. The image resolution is $1.0 \times 1.0 \times 1.0 \text{ mm}^3$. The scan matrix size is 512×512 , and the number of sliced images is between 35 and 98. The other dataset contains 30 individual images of segmented complete vertebral bodies with a resolution of $0.35 \times 0.35 \times 1 \text{ mm}^3$. The acquired slice image size is 512×512 . The number of slice images is 265 to 1050. In the dataset of 6 individual images, the number of network training, testing, and verification iterations is 3, 2, and 2, respectively, and the dataset containing 30 individuals is divided into 2 individual datasets for verification. Others are used for network training and cross testing. During each training iteration, the input trained image is elastically deformed by a density deformation field obtained by using a 3×3 grid control point and cubic B-spline interpolation. A new variant of the training dataset is derived. The primary purpose of this dataset is to verify the validity and reliability of the data expansion method.

The deep learning model used in this experiment was based on the Pytorch implementation and was trained on the

Titan-X GPU. The network architecture in the deep learning model is the network architecture proposed in Section 2 of this paper. The overfitting processing method of the deep learning model is the adaptive depth-based deep computing model proposed in Section 3. The initial learning rate was set to 0.001, and the learning rate was reduced by one tenth when the training progressed to 60 and 120 epochs. The training lasted for 200 epochs. In all training sessions, the model proposed in this paper is trained based on the stochastic gradient descent method, and the number of samples per batch is set to 64.

4.3.2. Segmentation Results and Analysis. The image segmentation algorithm proposed in this paper and other mainstream medical image segmentation algorithms are used to segment the experimental dataset, respectively. The partial image segmentation results are shown in Figure 5. It can be seen from Figure 5 that the results obtained by the proposed segmentation algorithm are the most satisfactory. The segmentation effects of the methods in [46, 47] have problems to varying degrees. For better comparison and analysis, Table 2 gives the specific segmentation results of the methods proposed in [46–50].

Table 2 shows that the segmentation effect of the medical image segmentation algorithm based on the optimized convolutional neural network with adaptive dropout depth calculation is better than that of the traditional machine learning image segmentation algorithms proposed in [45, 46] and [41, 47–49]. The segmentation effect of the proposed deep learning algorithm has advantages. Specifically, the traditional machine learning methods proposed in [45, 46] have the lowest Dice values after image segmentation, 0.8481 and 0.8570, respectively, and the Jaccard indexes are lower than 0.8, while the FP indexes are higher than 0.1. This shows that the machine learning method is one of the worst-case methods among the above-listed methods.

Krizhevsky et al. [41] proposed a deep learning model without optimized convolutional neural network model and improved dropout. The Dice index is higher than 0.92, the Jaccard index is higher than 0.86, and the FP index is lower than 0.07. It shows that the convolutional neural network model that has not been optimized has a significant improvement over traditional machine learning methods. This is mainly because the deep learning model can better train the experimental data and obtain a more reasonable and reliable image segmentation model. It confirms that the deep learning method is suitable for image segmentation. The image segmentation method of the deep learning method proposed in [48–50] has obtained Dice and Jaccard indexes both higher than 0.9, and some have reached more than 0.95, and the FP indexes are lower than 0.03. It shows that the segmentation effect of the deep learning method proposed in [48–50] is not only significantly improved compared with the traditional machine learning method but also superior to the unoptimized convolutional neural network model. It shows that improving and optimizing the convolutional neural network model can improve the image segmentation effect.

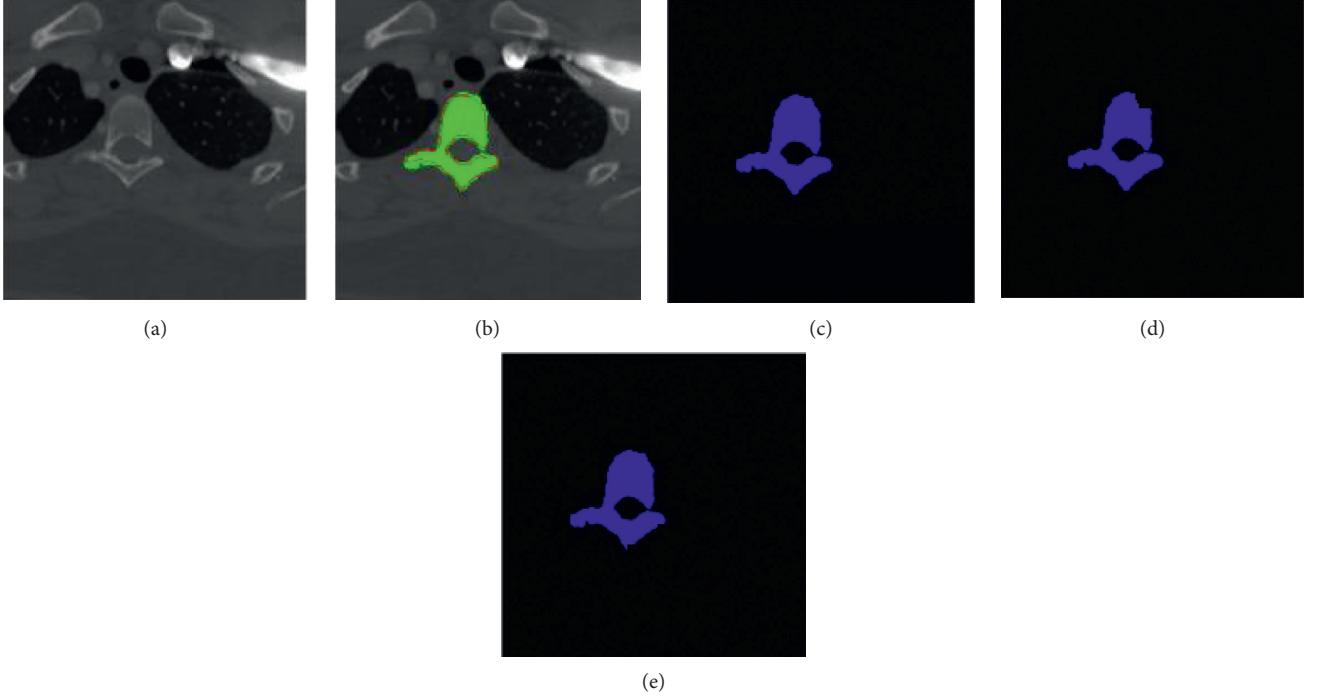


FIGURE 5: Partial CT image vertebral segmentation results. (a) Original image. (b) Standard result. (c) Ours. (d) Segmentation results of [45]. (e) Segmentation results of [46].

TABLE 2: Comparison of SpineWeb dataset segmentation results.

Segmentation method	Dice	Jaccard	FP
[46]	0.8481	0.7904	0.1270
[47]	0.8570	0.7986	0.1131
[41]	0.9206	0.8675	0.0649
[48]	0.9400	0.9108	0.0295
[49]	0.9496	0.9407	0.0200
[50]	0.9611	0.9526	0.0174
Ours	0.9932	0.9881	0.0010

The Dice and Jaccard indicators for image segmentation obtained by the method proposed in this paper are the highest among those of all the methods, and the FP index is the lowest among those of all the methods. This method is not only superior to traditional machine learning methods but also superior to other deep learning methods. This is mainly because the method proposed in this paper is better optimized than the deep learning methods proposed in [48–50]. The proposed method not only optimizes the network structure of the deep learning model but also solves the overfitting problem of the deep learning model. Through the above processing, the deep learning model has a stronger modeling ability and a nonlinear fitting energy. It can fully extract all kinds of characteristic information of lumbar CT images to obtain a better segmentation effect.

In short, the traditional machine learning segmentation algorithm has problems such as weak adaptive ability and poor segmentation effect in medical image segmentation. It shows that the traditional machine learning segmentation algorithm is difficult to adapt to the requirements of image

segmentation accuracy in the medical field. The depth learning image segmentation algorithm is significantly better than the traditional machine learning method in the above dataset. It validates the advantages of deep learning models in medical image segmentation. Among these deep learning models, the medical image segmentation algorithm based on optimized convolutional neural network-adaptive dropout depth calculation has better segmentation effect than other deep learning algorithms. This is because the deep learning model proposed in this paper better solves the problem of deep learning architecture optimization and overfitting.

5. Conclusion

To obtain a better medical image segmentation method based on deep learning, this paper first studies the network structure of the deep learning model and adds cross-layer connections to the traditional convolutional neural network. An optimized convolutional neural network model is established. In the optimized convolutional neural network model, two scale features can be used simultaneously for segmentation. At the same time, to solve the overfitting problem of the traditional deep learning model, this paper proposes an adaptive dropout model, which can improve the generalizability of the dropout model and better solve the overfitting problem, thereby addressing the segmentation problem of the corresponding medical image. Based on this idea, this paper proposes a medical image segmentation algorithm based on an optimized convolutional neural network with adaptive dropout depth calculation.

The experimental results for the ultrasonic tomographic image dataset and the SpineWeb dataset show that the

segmentation effect of the deep learning medical image segmentation method is the best. The Dice and Jaccard indicators of the segmented image are the highest, and the false positive index is the lowest. The data of these three indicators show that the segmentation method proposed in this paper obtains an optimal medical image segmentation effect for these two datasets. The method in this paper is not only superior to the traditional deep learning model but also superior to the deep learning model that is only optimized for the network without dropout optimization. The method proposed in this paper can obtain such an excellent segmentation effect primarily because, first, the deep learning method proposed in this paper solves the problem of network architecture optimization for the deep learning model. Second, the deep learning method proposed in this paper addresses the overfitting problem better than previous methods.

Data Availability

The data used to support the findings of this study are included within the paper.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (no. 61701188), China Postdoctoral Science Foundation (no. 2019M650512), and Beijing Intelligent Logistics System Collaborative Innovation Center (no. BILSCIC-2019KF-22).

References

- [1] M. Drozdzal, G. Chartrand, E. Vorontsov et al., “Learning normalized inputs for iterative estimation in medical image segmentation,” *Medical Image Analysis*, vol. 44, pp. 1–13, 2018.
- [2] Y. Xue, T. Xu, H. Zhang et al., “Segan: adversarial network with multi-scale l1 loss for medical image segmentation,” *Neuroinformatics*, vol. 16, no. 3–4, pp. 383–392, 2018.
- [3] S. Zhou, J. Wang, S. Zhang, Y. Liang, and Y. Gong, “Active contour model based on local and global intensity information for medical image segmentation,” *Neurocomputing*, vol. 186, pp. 107–118, 2016.
- [4] M. Vardhana, N. Arunkumar, S. Lasrado, E. Abdulhay, and G. Ramirez-Gonzalez, “Convolutional neural network for biomedical image segmentation with hardware acceleration,” *Cognitive Systems Research*, vol. 50, pp. 10–14, 2018.
- [5] G. Wang, M. A. Zuluaga, W. Li et al., “DeepIGeoS: a deep interactive geodesic framework for medical image segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 7, pp. 1559–1572, 2018.
- [6] A. Khadidos, V. Sanchez, and C.-T. Li, “Weighted level set evolution based on local edge features for medical image segmentation,” *IEEE Transactions on Image Processing*, vol. 26, no. 4, pp. 1979–1991, 2017.
- [7] B. Gaonkar, D. Hovda, N. Martin et al., “Deep learning in the small sample size setting: cascaded feed forward neural networks for medical image segmentation,” *International Society for Optics and Photonics*, vol. 97, 2016.
- [8] X. Wang, S. Han, Y. Chen et al., “Volumetric attention for 3D medical image segmentation and detection,” in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, Berlin, Germany, pp. 175–184, 2019.
- [9] A. Zhao, G. Balakrishnan, F. Durand et al., “Data augmentation using learned transformations for one-shot medical image segmentation,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8543–8553, New York, NY, USA, 2019.
- [10] R. Rodrigues, R. Braz, M. Pereira, J. Moutinho, and A. M. G. Pinheiro, “A two-step segmentation method for breast ultrasound masses based on multi-resolution analysis,” *Ultrasound in Medicine & Biology*, vol. 41, no. 6, pp. 1737–1748, 2015.
- [11] A. M. Anter and A. E. Hassenian, “CT liver tumor segmentation hybrid approach using neutrosophic sets, fast fuzzy c-means and adaptive watershed algorithm,” *Artificial Intelligence in Medicine*, vol. 97, pp. 105–117, 2019.
- [12] H. Verma, R. K. Agrawal, and A. Sharan, “An improved intuitionistic fuzzy c-means clustering algorithm incorporating local information for brain image segmentation,” *Applied Soft Computing*, vol. 46, pp. 543–557, 2016.
- [13] Q.-H. Huang, S.-Y. Lee, L.-Z. Liu, M.-H. Lu, L.-W. Jin, and A.-H. Li, “A robust graph-based segmentation method for breast tumors in ultrasound images,” *Ultrasonics*, vol. 52, no. 2, pp. 266–275, 2012.
- [14] S. Yu, S. Wu, L. Zhuang et al., “Efficient segmentation of a breast in B-mode ultrasound tomography using three-dimensional GrabCut (GC3D),” *Sensors*, vol. 17, no. 8, pp. 1827–1839, 2017.
- [15] Y. Guo, Y. Gao, Y. Shao et al., “Deformable segmentation of 3D MR prostate images via distributed discriminative dictionary and ensemble learning,” *Medical Physics*, vol. 41, no. 7, pp. 72–84, 2014.
- [16] G. E. Hinton and R. R. Salakhutdinov, “Reducing the dimensionality of data with neural networks,” *Science*, vol. 313, no. 5786, pp. 504–507, 2006.
- [17] Y. Bengio, P. Lamblin, D. Popovici et al., “Greedy layer-wise training of deep networks,” *Advances in Neural Information Processing Systems*, vol. 34, pp. 153–160, 2007.
- [18] O. Ronneberger, P. Fischer, and T. Brox, “U-net: convolutional networks for biomedical image segmentation,” in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, Berlin, Germany, pp. 234–241, 2015.
- [19] F. Milletari, N. Navab, and S. A. V. Ahmadi, “Fully convolutional neural networks for volumetric medical image segmentation,” in *Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV)*, IEEE, Berlin, Germany, pp. 565–571, 2016.
- [20] Y. Xu, Y. Wang, J. Yuan, Q. Cheng, X. Wang, and P. L. Carson, “Medical breast ultrasound image segmentation by machine learning,” *Ultrasonics*, vol. 91, pp. 1–9, 2019.
- [21] V. Kumar, J. M. Webb, A. Gregory et al., “Automated and real-time segmentation of suspicious breast masses using convolutional neural network,” *PloS One*, vol. 13, no. 5, pp. 1–18, 2018.
- [22] M. Ahmad, J. Yang, D. Ai et al., “Deep-stacked auto encoder for liver segmentation,” in *Proceedings of the Chinese Conference on Image and Graphics Technologies*, Springer, Singapore, pp. 243–251, 2017.

- [23] W. Tang, D. Zou, S. Yang et al., “DSL: automatic liver segmentation with faster R-CNN and deeplab,” in *Proceedings of the International Conference on Artificial Neural Networks*, Springer, Berlin, Germany, pp. 137–147, 2018.
- [24] P. Hu, F. Wu, J. Peng, P. Liang, and D. Kong, “Automatic 3D liver segmentation based on deep learning and globally optimized surface evolution,” *Physics in Medicine and Biology*, vol. 61, no. 24, pp. 8676–8698, 2016.
- [25] D. Yang, D. Xu, S. K. Zhou et al., “Automatic liver segmentation using an adversarial image-to-image network,” in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, Berlin, Germany, pp. 507–515, 2017.
- [26] X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, and P.-A. Heng, “H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes,” *IEEE Transactions on Medical Imaging*, vol. 37, no. 12, pp. 2663–2674, 2018.
- [27] C. Sun, S. Guo, H. Zhang et al., “Liver lesion segmentation in CT images with MK-FCN,” in *Proceedings of the 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, IEEE, Berlin, Germany, pp. 1794–1798, 2017.
- [28] Y. Guo, Y. Gao, and D. Shen, “Deformable MR prostate segmentation via deep feature learning and sparse patch matching,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 4, pp. 1077–1089, 2015.
- [29] D. Karimi, G. Samei, C. Kesch, G. Nir, and S. E. Salcudean, “Prostate segmentation in MRI using a convolutional neural network architecture and training strategy based on statistical shape models,” *International Journal of Computer Assisted Radiology and Surgery*, vol. 13, no. 8, pp. 1211–1219, 2018.
- [30] L. Yu, X. Yang, H. Chen et al., “Volumetric convnets with mixed residual connections for automated prostate segmentation from 3D MR images,” *Conference on Artificial Intelligence*, vol. 33, pp. 66–72, 2017.
- [31] N. Srivastava, G. Hinton, A. Krizhevsky et al., “Dropout: a simple way to prevent neural networks from overfitting,” *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [32] S. J. Rennie, P. L. Dognin, X. Cui et al., “Annealed dropout trained maxout networks for improved lvcsr,” in *Proceedings of the 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, Berlin, Germany, pp. 5181–5185, 2015.
- [33] H. Zheng, M. Chen, W. Liu et al., “Improving deep neural networks by using sparse dropout strategy,” in *Proceedings of the 2014 IEEE China Summit & International Conference on Signal and Information Processing (ChinaSIP)*, IEEE, Berlin, Germany, pp. 21–26, 2014.
- [34] G. Pandey and A. Dukkipati, “Variational methods for conditional multimodal deep learning,” in *Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN)*, IEEE, Berlin, Germany, pp. 308–315, 2017.
- [35] F. Bu, Z. Chen, P. Li, T. Tang, and Y. Zhang, “A high-order CFS algorithm for clustering big data,” *Mobile Information Systems*, Mobile Information Systems, vol. 2016, , pp. 1–8, 2016.
- [36] T. Hopp, M. Zapf, E. Kretzek et al., “3D ultrasound computer tomography: update from a clinical study,” *International Society for Optics and Photonics*, vol. 97, 2016.
- [37] T. Hopp, N. Duric, and N. V. Ruiter, “Image fusion of Ultrasound Computer Tomography volumes with X-ray mammograms using a biomechanical model based 2D/3D registration,” *Computerized Medical Imaging and Graphics*, vol. 40, pp. 170–181, 2015.
- [38] S. S. Choi, S. H. Cha, and C. C. Tappert, “A survey of binary similarity and distance measures,” *Journal of Systemics, Cybernetics and Informatics*, vol. 8, no. 1, pp. 43–48, 2010.
- [39] H. Pathak and V. Kulkarni, “Identification of ovarian mass through ultrasound images using machine learning techniques,” in *Proceedings of the 2015 IEEE International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*, IEEE, Berlin, Germany, pp. 137–140, 2015.
- [40] M. J. Gangeh, H. R. Tizhoosh, K. Wu et al., “Tumour ellipsification in ultrasound images for treatment prediction in breast cancer,” in *Proceedings of the 2017 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, IEEE, Berlin, Germany, pp. 41–44, 2017.
- [41] A. Krizhevsky, I. Sutskever, G. E. Hinton et al., “ImageNet classification with deep convolutional neural networks,” *Neural Information Processing Systems*, vol. 2012, pp. 1097–1105, 2012.
- [42] Y. Kim, C. Audigier, N. Ellens et al., “Low-cost ultrasound thermometry for HIFU therapy using CNN,” *IEEE*, vol. 2018, pp. 1–9, 2018.
- [43] Q. Zeng, G. Samei, D. Karimi et al., “Prostate segmentation in transrectal ultrasound using magnetic resonance imaging priors,” *International Journal of Computer Assisted Radiology and Surgery*, vol. 13, no. 6, pp. 749–757, 2018.
- [44] M. Fradi, P. Lasaygues, and M. Machhout, “Auto-organiser neural network application for ultrasound computed tomographic image classification,” in *Proceedings of the 2019 19th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)*, pp. 19–23, Berlin, Germany, 2019.
- [45] Y. Cai, M. Landis, D. T. Laidley, A. Kornecki, A. Lum, and S. Li, “Multi-modal vertebrae recognition using transformed deep convolution network,” *Computerized Medical Imaging and Graphics*, vol. 51, pp. 11–19, 2016.
- [46] S. Park, H. S. Lee, and J. Kim, “Seed growing for interactive image segmentation using SVM classification with geodesic distance,” *Electronics Letters*, vol. 53, no. 1, pp. 22–24, 2016.
- [47] X. Liang, L. Lin, W. Yang, P. Luo, J. Huang, and S. Yan, “Clothes Co-parsing via joint image segmentation and labeling with application to clothing retrieval,” *IEEE Transactions on Multimedia*, vol. 18, no. 6, pp. 1175–1186, 2016.
- [48] S. Bao and A. C. S. Chung, “Multi-scale structured CNN with label consistency for brain MR image segmentation,” *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 6, no. 1, pp. 113–117, 2018.
- [49] J. Dolz, K. Gopinath, J. Yuan et al., “HyperDense-Net: a hyper-densely connected CNN for multi-modal image segmentation,” *IEEE Transactions on Medical Imaging*, vol. 38, no. 5, pp. 1116–1126, 2018.
- [50] H. Kervadec, J. Dolz, M. Tang, E. Granger, Y. Boykov, and I. Ben Ayed, “Constrained-CNN losses for weakly supervised segmentation,” *Medical Image Analysis*, vol. 54, pp. 88–99, 2019.