J BONE ONCOL 45 (2024) 100593

Contents lists available at ScienceDirect

Journal of Bone Oncology

journal homepage: www.elsevier.com/locate/jbo

Pelvic bone tumor segmentation fusion algorithm based on fully convolutional neural network and conditional random field

Shiqiang Wu [a,b,*,1], Zhanlong Ke [b,1], Liquan Cai [a], Liangming Wang [a], XiaoLu Zhang [a], Qingfeng Ke [a], Yuguang Ye [c,d,e,*]

[a] Department of Orthopedics, The Second Affiliated Hospital of Fujian Medical University, Quanzhou, Fujian 362000, China
[b] Department of Orthopedics, The Second Clinical College of Fujian Medical University, Quanzhou, Fujian 362000, China
[c] Faculty of Mathematics and Computer Science, Quanzhou Normal University, Quanzhou 362000, China
[d] Fujian Provincial Key Laboratory of Data Intensive Computing, Quanzhou 362000, China
[e] Key Laboratory of Intelligent Computing and Information Processing, Fujian Province University, Quanzhou 362000, China

HIGHLIGHTS

- Current machine learning algorithms for pelvic bone tumor image segmentation have limited accuracy.
- Our proposed algorithm combines a fully convolutional neural network and a conditional random field to achieve more accurate segmentation of pelvic bone tumor images.
- FCNN-4s is used to improve the precision and convergence speed of pelvic bone tumor segmentation.
- FCNN-4s adopts operations like Crop and Fuse, padding, ReLU activation, and SoftMax loss with optimized hyperparameters for better performance.
- Our algorithm demonstrated an improvement of 6.69% in terms of the Dice coefficient compared to other algorithms, with an average enhancement of 9.33%

ARTICLE INFO

Keywords:
Pelvis bone tumor
Image segmentation
Fully convolutional neural network
Conditional random fields
Computed tomography

ABSTRACT

Background and objective: Pelvic bone tumors represent a harmful orthopedic condition, encompassing both benign and malignant forms. Addressing the issue of limited accuracy in current machine learning algorithms for bone tumor image segmentation, we have developed an enhanced bone tumor image segmentation algorithm. This algorithm is built upon an improved full convolutional neural network, incorporating both the fully convolutional neural network (FCNN-4s) and a conditional random field (CRF) to achieve more precise segmentation.

Methodology: The enhanced fully convolutional neural network (FCNN-4s) was employed to conduct initial segmentation on preprocessed images. Following each convolutional layer, batch normalization layers were introduced to expedite network training convergence and enhance the accuracy of the trained model. Subsequently, a fully connected conditional random field (CRF) was integrated to fine-tune the segmentation results, refining the boundaries of pelvic bone tumors and achieving high-quality segmentation.

Results: The experimental outcomes demonstrate a significant enhancement in segmentation accuracy and stability when compared to the conventional convolutional neural network bone tumor image segmentation algorithm. The algorithm achieves an average Dice coefficient of 93.31 %, indicating superior performance in real-time operations.

Conclusion: In contrast to the conventional convolutional neural network segmentation algorithm, the algorithm presented in this paper boasts a more intricate structure, proficiently addressing issues of over-segmentation and under-segmentation in pelvic bone tumor segmentation. This segmentation model exhibits superior real-time performance, robust stability, and is capable of achieving heightened segmentation accuracy.

* Corresponding authors at: Department of Orthopedics, The Second Affiliated Hospital of Fujian Medical University, Quanzhou, Fujian 362000, China (S. Wu). Faculty of Mathematics and Computer Science, Quanzhou Normal University, Quanzhou 362000, China (Y. Ye).
E-mail addresses: wushiqiang@fjmu.edu.cn (S. Wu), chinaye@qztc.edu.cn (Y. Ye).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.jbo.2024.100593
Received 7 October 2023; Received in revised form 24 January 2024; Accepted 8 February 2024
Available online 28 February 2024

1. Introduction

Pelvic bone tumors account for about 4 % of bone tumors in the whole body, and the morphology of the lesions is complex and changeable, which can involve one or more pelvic divisions of the pelvis, and there is significant heterogeneity [1]. Pelvic tumors, whether benign or malignant, can affect the local bony structure, as an axial weight-bearing bone, there is a potential risk of pathological fracture. Malignant tumors can also occur local destructive growth, further invade adjacent organs, and are prone to recurrence and metastasis, resulting in limb dysfunction of patients. If it cannot be fully, timely and standardized treatment, it will endanger the lives of patients.

The determination of tumor boundary is the focus of imaging preacademic planning. The accurate determination of the image boundary of pelvic tumor requires some practice and experience, and the images of different modes have their own advantages: CT images have clear bone structure, can clearly show the characteristic changes such as bone destruction, periosteal reaction, tumor bone and calcification, and have high spatial resolution, which is convenient for the three-dimensional modeling of bony anatomical structure. MRI soft tissue resolution is very high, so it is regarded as the "gold standard" for tumor boundary determination. Combined with CT and MRI angiography, the anatomical relationship between the focus area and the surrounding vessels can be observed. According to the signal characteristics and pathological tissue characteristics of MRI images, tumors can be roughly divided into tumor central area, tumor marginal zone, surrounding edema area and surrounding transitional zone, and different MRI sequences have different sensitivity to different regions, and the manifestations of soft tissue masses inside and around the bone marrow of the tumor are different.

In recent years, due to the rapid development of computer technology, computer vision, image processing, and pattern recognition, digital image segmentation technology has played an increasingly important role in quantitative analysis of medical images. Digital image segmentation is the process of dividing an image into multiple regions based on specific attributes (such as color, texture, density, etc.). Accurate image segmentation is the foundation of quantitative analysis of medical images. Despite the relatively recent emergence of digital image processing, research on image segmentation technology has made considerable progress and success. Zhou et al. [2] proposed a method for recognizing precancerous lesions and cancers in gastric cancer based on the fusion of shallow and deep features in endoscopic images. However, to date, there is still no segmentation methods are suitable for a particular class of images.

Due to the similar X-ray absorption coefficients between pelvis tumors and surrounding tissues, as well as the effects of X-ray scattering, uneven light intensity, and different irradiation angles, the edges of bone tumors in X-ray images appear blurry, and the distribution of tumor and background features overlap. Therefore, it is difficult to obtain satisfactory segmentation results using various global thresholding techniques reviewed in the literature [3,4] alone. It is necessary to fully consider the local characteristics of the images.

With the rapid development of computer technology, computer vision, image processing, and pattern recognition, digital image segmentation plays an increasingly important role in quantitative analysis of medical images. Among them, MRI and computed tomography (CT) are widely used imaging techniques for detecting abnormalities in tumor shape, size, or location, which help in detecting tumors. Zhao et al. [5] proposed a novel GAN that utilizes its synthesized CT images to visually resemble the reference CT (RCT) images and achieve desirable results on local mismatched tissues. In quantitative evaluation, it outperforms other compared methods. In addition, MRI-guided radiotherapy is a current research hotspot in radiotherapy. Brown et al. [6] used radiation features extracted from MR images to establish a model to distinguish radiation necrosis from tumor progression in brain metastases after gamma knife radiosurgery.

2. Materials and methods

2.1. Framework for diagnostic medical image analysis using deep learning

2.1.1. Enhanced fully convolutional neural network

In 1998, LeCun et al. [7] first applied CNN to image recognition, and in 2012, the AlexNet network structure proposed by Krizhevsky et al. [8] made significant breakthroughs in the field of CNN object classification. In 2015, Long et al. [9] proposed a pixel-level image semantic segmentation network called Fully Convolutional Neural Network (FCNN), which improved upon AlexNet by performing end-to-end pixel-wise classification and achieving image segmentation tasks. FCNN converted the two fully connected layers of AlexNet into convolutional layers, allowing the network's final output to remain a two-dimensional matrix, thus preserving spatial information between pixels and facilitating feature extraction. To make the output of the network the same size as the input, the original FCNN simply upsampled the output of the last convolutional layer to the same size as the input image, but this operation only utilized information from the 5th pooling layer, resulting in a coarser target feature that made it difficult to achieve precise segmentation.

This algorithm will use feature information from different pooling layers to achieve multi-level feature fusion. After upsampling the output of the last convolutional layer, the first fusion will be performed with the feature vector matrix from the fourth pooling layer, and then the resulting feature fusion matrix will be upsampled and fused with the feature vector matrix from the third pooling layer. Then, the fusion feature matrix will be upsampled and fused with the feature vector matrix from the second pooling layer, and finally, the feature matrix from the third fusion will be upsampled to obtain a feature matrix the same size as the original image, resulting in a better feature information matrix. At the same time, the algorithm adds the batch normalization (BN) layer proposed by Ioffe et al. [10] after each convolutional layer to speed up network training and improve segmentation accuracy. The network structure with multi-level feature fusion is shown in Fig. 1, where Conv represents a convolutional layer that can extract various features of the image such as edges and positions, Pool represents a pooling layer that can achieve feature dimensionality reduction and preserve the main features extracted by the convolutional layer, BN represents a batch normalization layer that ensures that the weight distribution of the network after convolution does not change significantly, Up represents an upsampling layer that mainly increases the size of the feature matrix through deconvolution, Crop_Fuse performs cropping and fusion operations on the feature matrix, and Pixelwise Prediction is the pixel-wise classification prediction layer that achieves image segmentation by classifying each pixel.

2.1.2. Conditional random field

Conditional Random Field (CRF) based image semantic segmentation is one of the classic probabilistic graphical segmentation algorithms, which can remove noise in the segmentation result and enhance the boundary segmentation of the image. This algorithm integrates the fully connected CRF model proposed by Zheng et al. [11] into the improved FCNN network for post-processing bone tumor boundary segmentation, thereby improving the segmentation accuracy of bone tumor images.

In the fully connected CRF model, the energy function that assigns pixels to their respective labels can be represented as:

$$E(x) = \sum_i \varphi_u(x_i) + \sum_{i \neq j} \varphi_p(x_i, x_j) \quad (1)$$

In Eq. (1), $E(x)$ represents the total energy of assigning pixels to their respective labels; $\varphi_u(x_i)$ is a unary energy potential function that represents the energy of assigning pixel i to label x_i without considering the relationship between pixels; $\varphi_p(x_i, x_j)$ is a pairwise energy potential function that represents the energy of assigning pixels i and j to labels x_i

and x_j based on the difference in grayscale values and spatial position distances between pixels, and describes the relationships between pairwise pixels, such that similar pixels are assigned the same label. The unary energy potential function $\varphi_u(x_i)$ can be obtained from the initial coarse segmentation result of FCNN, and the pairwise energy potential function $\varphi_p(x_i, x_j)$ can be represented as:

$$\varphi_p(x_i, x_j) = \mu(x_i, x_j)\sum_{m=1}^{M} W^{(m)} k_G^{(m)}(f_i, f_j) \tag{2}$$

In Eq. (2), $\mu(x_i, x_j)$ is the label compatibility matrix, which represents the penalty for assigning different labels to different pixels. When $x_i \neq x_j$, the penalty $\mu(x_i, x_j) = 1$; $k_G^{(m)}(f_i, f_j)$ is the Gaussian filter kernel, where f_i, and f_j are the feature vectors of the filter, determined by the spatial position between pixels and their grayscale values, and m is the number of filters; $W^{(m)}$ is the weight of each filter.

In the CRF model, a bilateral Gaussian filter is used to assign pixels that are close in space and have similar grayscale values to the same label; a spatially smooth Gaussian filter is used to remove isolated small regions in the bone tumor segmentation result. Therefore, Eq. (2) can be further represented as:

$$\varphi_p(x_i, x_j) = \mu(x_i, x_j)\left[W^{(1)} exp\left(-\frac{|P_i - P_j|}{2\theta_\alpha^2} - \frac{|R_i - R_j|}{2\theta_\beta^2}\right) + W^{(2)} exp\left(-\frac{|P_i - P_j|}{2\theta_\gamma^2}\right)\right] \tag{3}$$

In Eq. (3), R_i and R_j represent the grayscale feature value vectors of pixels i and j; P_i and P_j represent the spatial position relationship feature vectors of pixels i and j; θ_α,θ_β and θ_γ represent the weights of the grayscale, position, and other factors on the potential functions of the pixels. Zheng et al. [9] have conducted a large number of experiments and found that when the bilateral filter weight $W^{(1)} = 5$, the control parameters $\theta_\alpha = 160$, $\theta_\beta = 3$, the smoothing filter weight $W^{(2)} = 3$, and the control parameter $\theta_\gamma = 5$, the boundary of the image can be better segmented.

2.2. The methodologies employed in this paper

In this paper, the algorithm flow of pelvis bone tumor image segmentation using the improved full convolutional neural network is shown in Fig. 2. Firstly, the collected images are preprocessed. Then, part of the processed data is used as the training set and the rest as the test set. This training set is used to train FCNN and CRF fusion algorithm. Finally, the model is tested on the test set, and the segmentation results are evaluated using performance indicators.

2.2.1. Image pre-processing

Each pixel in the original CT pelvis bone tumor image is stored in the computer as 16 bits, but in digital image processing, 8-bit images are commonly used. Therefore, in this study, the CT images were first normalized in terms of gray level, with the gray values of each pixel uniformly compressed to the range of 0–255.

2.2.2. Enhanced FCNN-4s initial segmentation algorithm

This paper introduces an improved FCNN algorithm, called FCNN-4s, which aims to improve the accuracy and convergence speed of pelvis bone tumor segmentation. In order to obtain good tumor features, this paper proposes a fine feature fusion model and adds a BN layer to the network (as shown in Fig. 3). The FCNN-4s algorithm includes a data layer, convolution layer, pooling layer, activation function layer, upsampling layer, and output probability map layer. The feature fusion operation (Fuse) is used to fuse the features of high-dimensional feature matrices and low-dimensional feature matrices, and the data of the corresponding dimensions of the two matrices are added together under the premise of the same dimension and size. If the dimensions and sizes of the two target matrices are not the same, the FCNN-4s algorithm will perform a matrix cropping operation (Crop) to make their dimensions and sizes the same, so that the network can smoothly learn the features. In the matrix cropping operation, if the dimensions are different, the low-dimensional feature matrix is convolved to make the two matrices have the same dimensions; if the sizes are different, the same size as the high-dimensional feature matrix is extracted symmetrically from the center of the low-dimensional feature matrix. These operations are all designed to improve the performance of the network and achieve better pelvis bone tumor segmentation results. To prevent the feature information matrix from becoming too small after pooling, this paper performs a padding operation on the original pelvis bone tumor image (240 × 240) with a padding size of 100. The output size indicates the dimension and size of the output feature matrix. The algorithm performs three fusion operations on the feature matrices of different dimensions to obtain more refined bone tumor features. Meanwhile, the BN layer can keep the weight distribution of the parameters during each iteration from changing significantly, thereby accelerating the convergence speed.

2.2.3. Refined fusion algorithm for FCNN-4s and CRF fine segmentation

This paper describes the process of determining whether each pixel in an CT pelvis bone tumor image is a tumor point, using a binary classification approach where the label "0" represents non-tumor pixels and "1" represents tumor pixels. After obtaining two initial probability maps from FCNN-4s, the energy function is initialized to obtain the original probability values for each pixel. The following steps are used to

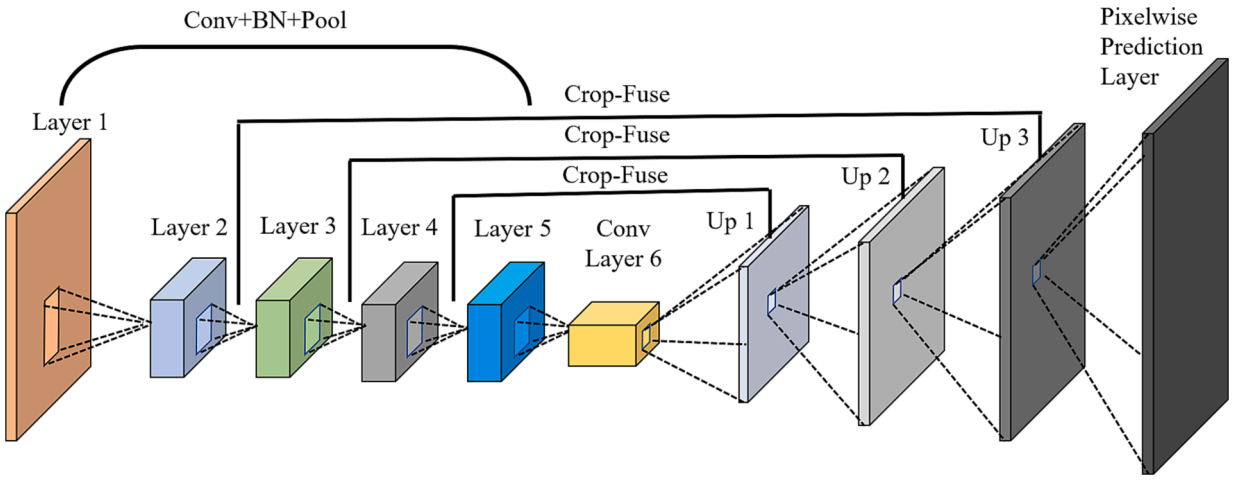

Fig. 1. The network structure chart of the multilevel features information fusion.

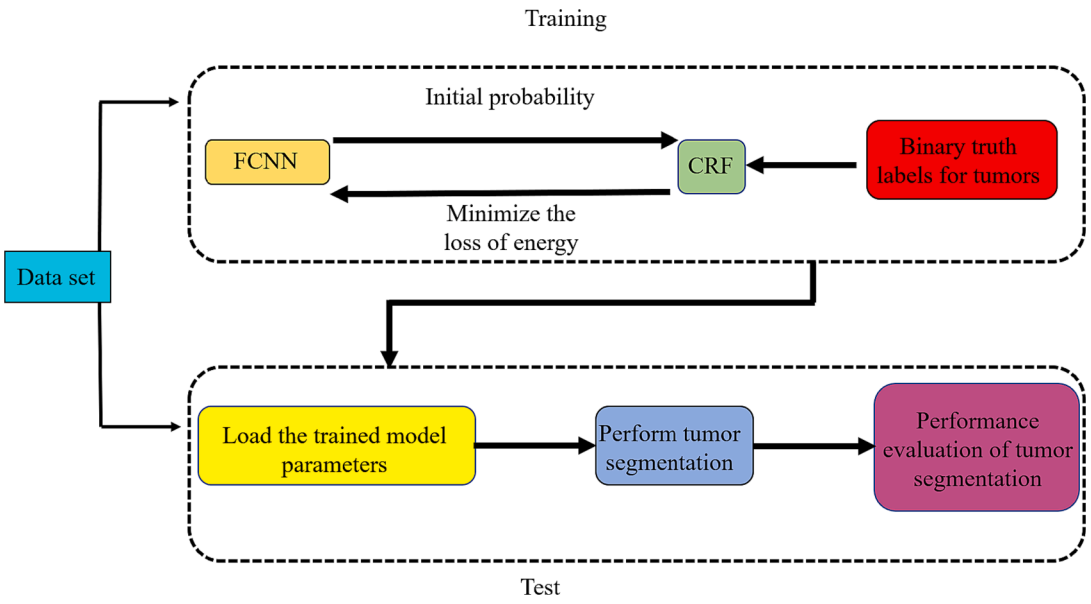

Fig. 2. The flow chart of algorithm.

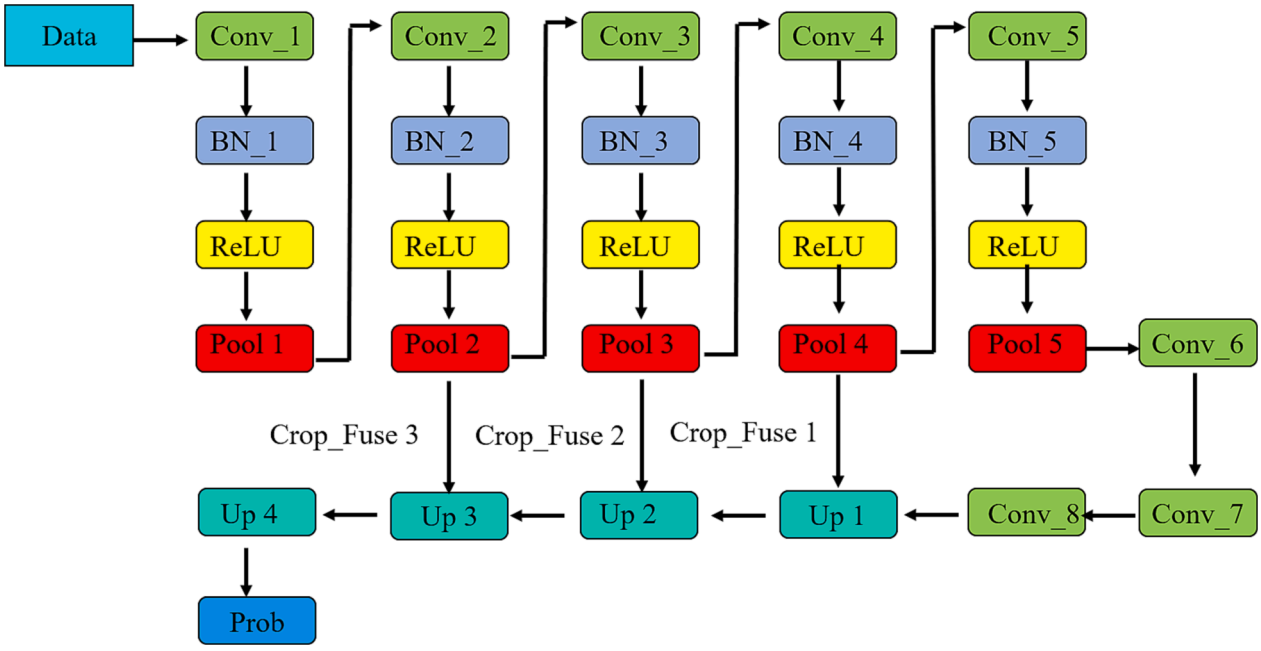

Fig. 3. The network structure chart of the improved FCNN-4s.

calculate the CRF model and iteratively adjust the two-class probability maps predicted by FCNN-4s using a fusion algorithm.

Step 1: Involving filtering the probability maps for both classes using a Gaussian filter $k_G^{(m)}\left(f_i, f_j\right)$ to obtain two filtered results $\widetilde{Q}^{(m)}(l)$ for labels "0″ and "1″, respectively.

Step 2: Assigning weights $W^{(m)}$ to $\widetilde{Q}^{(m)}(l)$ and calculates the pairwise energy potential function $\varphi_p(x_i, l)$ for each class probability map based on the label compatibility matrix $\mu^{(m)}(x_i, l)$.

Step 3: Calculating the unary energy potential function $\varphi_u(x_i)$ for the output of the FCNN-4s network, and then integrates it with the pairwise energy potential function $\varphi_p(x_i, l)$ from Step 2 to obtain the overall energy function $\widehat{Q}_i(x_i)$.

Step 4: Normalizing $\widehat{Q}_i(x_i)$ to obtain the probability values for each pixel's label, denoted as $Q_i(x_i)$. and selects the label with the highest probability for each pixel. The algorithm loop continues until the probability values for each pixel's label converge to 90 % or more, at which point the algorithm stops. If the convergence threshold is not met, the pixel loss is back-propagated to the FCNN-4s algorithm for further learning and parameter adjustment.

The pseudo-code for the algorithm is shown below, which results in a fusion network structure of CRF and FCNN-4s. The end-to-end calculation of the loss and parameter updates for each pixel can be performed using the back-propagation algorithm.

Begin

$Q_i(x_i) \leftarrow \frac{1}{Z_i} exp(-\varphi_u(x_i))$ for all x_i, Z is the normalization factor.

While not converged $Q_i(x_i) \geq 90\%$ do

(*continued on next page*)

(*continued*)

$\widetilde{Q}_i^{(m)}(l) \leftarrow \sum_{j \neq i} k_G^{(m)}\left(f_i, f_j\right) Q_j(l)$ *for* $m = 2$;

$\varphi_p(x_i, l) \leftarrow \sum_{l \in L} \mu^{(m)}(x_i, l) \sum_{m=1}^{2} W^{(m)} \widetilde{Q}_i^{(m)}(l)$;

$\widehat{Q}_i(x_i) \leftarrow \varphi_u(x_i) + \varphi_p(x_i, l)$;

$Q_i(x_i) \leftarrow \frac{1}{Z} exp(-\widehat{Q}_i(x_i))$;

end while

End

2.3. Segmentation using U-Net

Other types of deep learning algorithms such as U-Net [12–14] may also be implemented for tumor segmentation. U-Net is a popular deep learning algorithm that has shown promising results in various medical image segmentation tasks, including tumor segmentation. It is a type of convolutional neural network that is designed to learn from both low-level and high-level features of the image, allowing for more accurate and robust segmentation. U-Net has been used in several studies for tumor segmentation, and its performance has been shown to be comparable or even superior to other segmentation methods. Its ability to handle complex and irregular shapes of tumors makes it a suitable choice for many medical image analysis applications.

Compared to typical CNN (Convolutional Neural Network) architectures, this network requires fewer training images and produces more accurate results. The U-Net processes input training images through a patch-based approach, equivalent to data augmentation. This characteristic reduces the demand for a large training dataset [15], enabling it to achieve excellent segmentation outcomes even with a small number of input samples. As illustrated in Fig. 4, the U-Net structure consists of 5 layers, primarily composed of convolution, pooling, and upsampling operations. The gray arrow operations in the diagram indicate the duplication and cropping of feature maps from the downsampling layers, which are then connected to corresponding upsampling layers.

2.4. Training procedure for the model

During the training stage, the mature Caffe deep learning framework was used to learn the model parameters on the training set. Considering that Sigmoid and Tanh activation functions easily cause gradient disappearance, ReLU was chosen as the activation function for the network, and the conventional SoftMax classification loss function was used.

Since the selection of learning rate is crucial to the network and determines whether the network can converge and to what degree, it cannot be too large or too small. In this paper, the convergence of the network under different learning rates was compared using the original training set. According to the experimental results, when the network selects relatively large learning rates such as 10e−08 and 10e−09, the training loss cannot converge. When the learning rate is 10e−10, the training loss can converge and the convergence is better than that of smaller learning rates such as 10e−11. Therefore, a learning rate of 10e−10 was ultimately adopted.

The weight decay coefficient has a certain impact on the degree of overfitting of the network. This paper compared the training accuracy of the network under different weight decay coefficients. The experimental results showed that when the weight decay coefficient is too small, it has not played a role in weight decay and the training accuracy of the network is poor. When the weight decay coefficient is too large, the network will be overfitting to some extent and the training accuracy will decrease. Therefore, a weight decay coefficient of 0.0005 was ultimately adopted.

To verify the effectiveness of adding BN layers, this paper compared the training loss and accuracy of the network before and after adding BN layers using the original training set. The network with BN layers had a loss of about 5000 after 8000 iterations, and the network had basically converged after 1.2×10^5 iterations. In contrast, the network without BN layers had a loss of about 5000 after 2.0×10^4 iterations, and the network converged after nearly 1.6×10^5 iterations, which fully demonstrates that BN layers can speed up the convergence of the network and reduce the loss value. Adding BN layers also improves the accuracy of the trained model to some extent.

2.5. Dataset and image pre-processing

The images used in the experiments were collected from The Second Affiliated Hospital of Fujian Medical University. The size of each CT image is 250 × 250 pixels, and the ground truth segmentation to be superimposed onto the CT image.are manually annotated by five experts.

In this experiment, 1500 CT images of pelvis bone tumors were randomly selected as the training set. The image preprocessing techniques and data augmentation techniques described in Section 2.2.1 were used to obtain a total of 48,000 training images with a size of 240 × 240 pixels. This paper mainly performed data augmentation on the pelvis bone tumor images, including horizontal and vertical flipping, as well as counterclockwise rotation by 90°, 180°, and 270°, to improve the accuracy of the trained model.

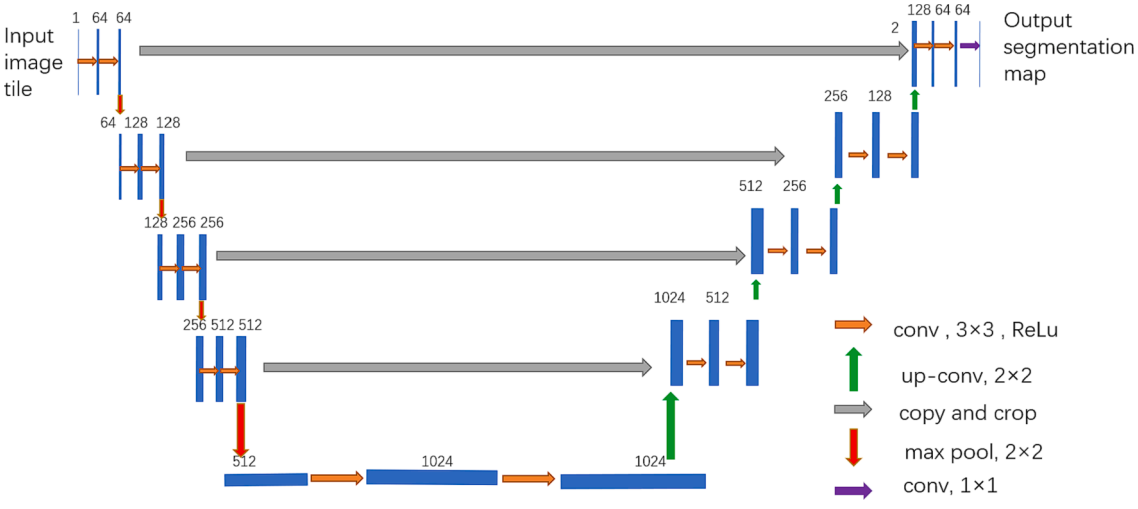

Fig. 4. U-Net Network Architecture.

3. Results and discussion

3.1. Evaluation criteria

This paper evaluates the segmentation results of bone tumors using indicators such as the Dice Similarity Coefficient (DSC), sensitivity, and positive predictive value (PPV). The DSC represents the similarity between the experimental segmentation results and the labels, sensitivity represents the ratio of correctly segmented tumor points to the true tumor points, and PPV represents the ratio of correctly segmented tumor points to the total segmented tumor points. The equations are as follows:

$$Dice = \frac{|P \wedge T|}{(|P| + |T|)/2} \tag{4}$$

$$Sensitivity = \frac{|P \wedge T|}{|T|} \tag{5}$$

$$Positive = \frac{|P \wedge T|}{|P|} \tag{6}$$

In the equations, P represents the segmentation results obtained using the algorithm proposed in this paper, while T represents the labels generated by experts for the bone tumors.

3.2. Training results

After correctly selecting the network architecture and parameters, to verify the impact of augmented datasets on model accuracy, this study compared the pelvis bone tumor classification accuracy across different datasets. As the dataset augmentation progressed, there was a noticeable improvement in training accuracy. Both the accuracy values and loss values during training are depicted in Fig. 5. This observation suggests that increasing the training samples can mitigate network overfitting and enhance the model's generalization capability. This ensures that the trained model can achieve robust pelvis bone tumor segmentation results on the test set.

3.3. Comparison between initial and refined segmentation

Our proposed pelvis bone tumor segmentation model operates in two distinct stages. In the initial stage, we employ the enhanced FCNN-4s algorithm for coarse segmentation. In the subsequent stage, we integrate the CRF model to perform fine segmentation, thereby enhancing the precision of our segmentation. Fig. 6 provides a visual representation of our algorithm's performance in both coarse and fine segmentation. Notably, our algorithm accurately identifies and segments the pelvis bone tumor, meticulously outlining its boundaries with minimal deviation when compared to the ground truth segmentation. This comparison, as depicted by the lines superimposed on the CT image in Figs. 6 and 7, illustrates our algorithm's exceptional accuracy.

To further quantify our results, we employed three key metrics for a comprehensive analysis, as summarized in Table 1. Our fine segmentation model significantly improved segmentation accuracy across various metrics, including the similarity index, sensitivity, and positive predictive value. Notably, our fine segmentation model achieved a 6.20 % improvement in positive predictive value (PPV).

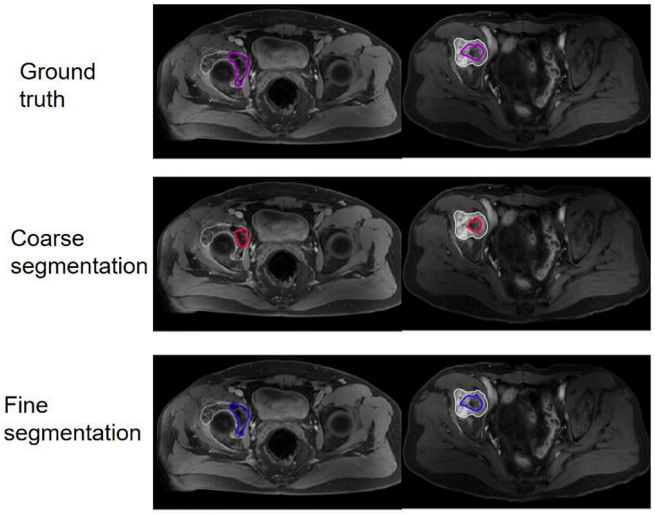

Fig. 6. The results of coarse segmentation based on FCNN-4s and fine segmentation based on FCNN-4s + CRF are compared.

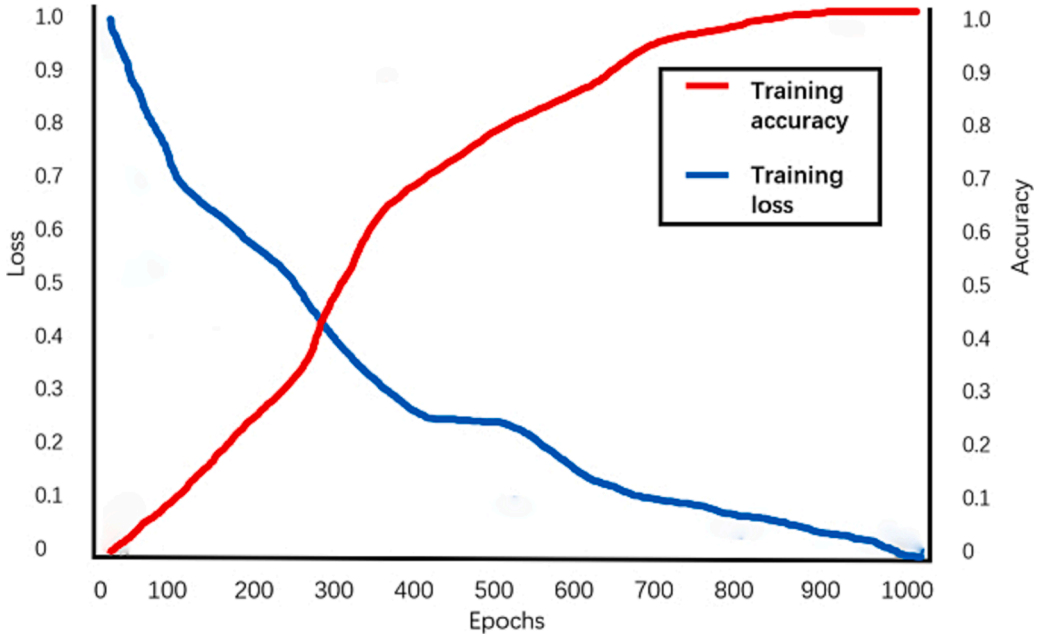

Fig. 5. Training model loss/accuracy curve graph.

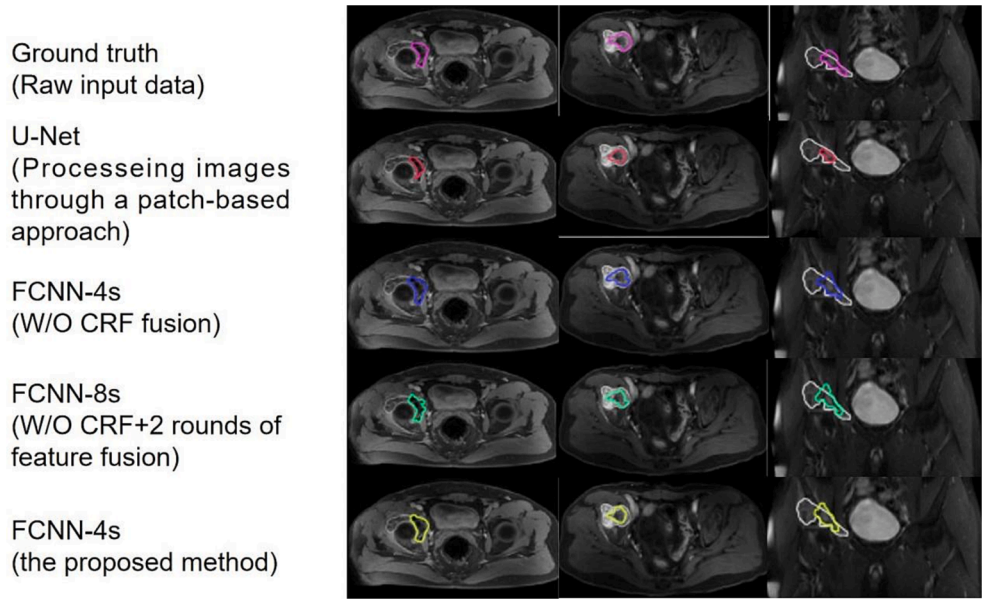

Fig. 7. Comparison of the segmentation results of based on four machine learning methods using fully Convolutional neural networks and conditional random field.

Table 1
Segmentation performance evaluation of different segmentation algorithms.

Different segmentation algorithm	Dice	Sensitivity	Positive
Improved FCNN-4s coarse segmentation algorithm	0.8831	0.8487	0.8589
The proposed method fine segmentation	0.9098	0.9401	0.9157

3.4. Comparison with other existing segmentation methods

In order to validate the superiority of our enhanced algorithm, we conducted comparisons with alternative segmentation methods. This included a simple segmentation algorithm, FCNN-8s, which incorporates only two rounds of feature fusion, as well as FCNN-4s without CRF fusion. Additionally, we compared our method to traditional CNN algorithms such as U-Net. Fig. 7 provides a visual comparison of the segmentation results achieved by these methods. Notably, the U-Net algorithm exhibited pronounced over-segmentation, resulting in unclear boundaries of pelvis bone tumors and numerous isolated points. The FCNN-8s algorithm, due to its limited feature fusion, yielded less refined tumor boundaries and reduced segmentation accuracy. In contrast, our FCNN-4s model, enriched with feature fusion and an end-to-end algorithmic structure that incorporates CRF fusion, consistently delivered satisfactory results. It effectively addressed issues associated with over-segmentation and under-segmentation of pelvis bone tumors.

Furthermore, as highlighted in Table 2, our proposed algorithm consistently demonstrated higher segmentation accuracy in comparison to alternative methods. Notably, when compared to the U-Net algorithm, our method achieved a 6.69 % improvement in Dice coefficient. Moreover, when compared to FCNN without CRF fusion, our algorithm showcased an average enhancement of 9.33 %. Additionally, in terms of the average time required for segmenting pelvis bone tumor images, our algorithm exhibited exceptional real-time performance during the prediction phase, completing the segmentation process within an average of 1 s.

Table 2
The segmentation performance evaluation of five algorithms.

Different segmentation algorithm	Dice	Sensitivity	Positive	The average time to complete an image segmentation
U-Net	0.8544	0.8123	0.8269	2.8324
FCNN-4s	0.8684	0.8636	0.8663	0.9779
FCNN-8s	0.8020	0.8283	0.8693	0.8673
The proposed method	0.9100	0.8982	0.9257	1.0062

Looking ahead, our future work may involve simulating bone cancer growth using discrete element methods [16] on high-performance computing platforms. Additionally, we can explore visually analyzing cancer spread and treatment using state-of-the-art graphical and meshing algorithms [17,18]. These computational simulations have the potential to enhance pelvis bone cancer diagnosis, thereby advancing orthopedic diagnostics. It is worth noting that deep learning will continue to play a crucial role in computer-aided diagnostics, offering promising technological advancements [19,20]. The future of medical image processing, coupled with advanced computational algorithms [21–24], holds great promise and is poised to deliver clinically viable diagnostic techniques for evaluating complex orthopedic conditions.

4. Conclusion

In this study, we present an enhanced pelvis bone tumor image segmentation algorithm designed to overcome the limitations of conventional Convolutional Neural Network (CNN) approaches, which often suffer from elevated computational complexity and diminished accuracy. Our novel algorithm leverages an improved Fully Convolutional Neural Network (FCNN-4s) for the preliminary segmentation of bone tumor pixels. Additionally, it incorporates a probabilistic graphical model that capitalizes on label correlations within the images and employs Conditional Random Field (CRF) to create a trainable, end-to-end segmentation framework.

Compared to conventional CNN-based segmentation methods, our algorithm, denoted as FCNN-4s + CRF, offers a more refined structural design. This refinement effectively addresses the issues of over-segmentation and under-segmentation encountered in pelvis bone tumor segmentation. Furthermore, it enhances segmentation prediction in terms of real-time performance and stability while achieving superior segmentation accuracy.

Our experimental findings underscore the superiority of the proposed algorithm, surpassing the performance of alternative methods. Notably, it excels in real-time performance, requiring an average of just 1 s to complete the segmentation of a pelvis bone tumor image.

Ethical approval

All human subjects in this study have given their written consent for the participation of our research.

CRediT authorship contribution statement

Shiqiang Wu: Writing – original draft, Data curation, Conceptualization. **Zhanlong Ke:** Writing – original draft, Formal analysis. **Liquan Cai:** Resources. **Liangming Wang:** Software. **XiaoLu Zhang:** Supervision. **Qingfeng Ke:** Visualization. **Yuguang Ye:** Writing – review & editing, Software, Methodology, Data curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This research is supported by Natural Science Foundation Project of Fujian Province (No. 2021J01274).

References

[1] (a) T. Damron, Dahlin's Bone Tumors: General Aspects and Data on 10,165 Cases, 6th ed. J. Bone Joint Surg. Am. 92(12) (2010) e17(11). (b) H. Zhou, Z. Liu, T. Li, Y. Chen, W. Huang, Z. Zhang, Classification of precancerous lesions based on fusion of multiple hierarchical features, Comput. Methods Programs Biomed. 229 (2022) 107301.

[2] H. Zhou, Z. Liu, T. Li, Y. Chen, W. Huang, Z. Zhang, Classification of precancerous lesions based on fusion of multiple hierarchical features, Comput. Methods Programs Biomed. 229 (2022) 107301.

[3] P.K. Sahoo, S. Soltani, A.K.C. Wong, A survey of thresholding techniques, Comput. Vision Graph. Image Process. 41 (3) (1988) 233–260.

[4] N.R. Pal, S.K. Pal, A review of image segmentation techniques, Pattern Recogn. 26 (9) (1993) 1277–1294.

[5] B. Zhao, T. Cheng, X. Zhang, J. Wang, H. Zhu, R. Zhao, D. Li, Z. Zhang, G. Yu, CT synthesis from MR in the pelvic area using residual transformer conditional GAN, Comput. Med. Imaging Graph. 103 (2023) 102150.

[6] Z. Zhang, J. Yang, A. Ho, W. Jiang, J. Logan, X. Wang, P.D. Brown, S.L. McGovern, N. Guha-Thakurta, S.D. Ferguson, X. Fave, L. Zhang, D. Mackin, L.E. Court, J. Li, A predictive model for distinguishing radiation necrosis from tumour progression after gamma knife radiosurgery based on radiomic features from MR images, Eur. Radiol. 28 (6) (2018) 2255–2263.

[7] Y. Lecun, L. Bottou, Y. Bengio, et al., Gradient-based learning applied to document recognition, Proc. IEEE 86 (11) (1998) 2278–2324.

[8] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, in: International Conference on Neural Information Processing Systems, Curran Associates Inc., 2012, pp. 1097–1105.

[9] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Computer Vision and Pattern Recognition, IEEE, 2015, pp. 3431–3440.

[10] S. Ioffe, C. Szegedy, Batch normalization: accelerating deep network training by reducing internal covariate shift, in: International Conference on Machine Learning, JMLR.org, 2015, pp. 448–456.

[11] S. Zheng, S. Jayasumana, B. Romera-Paredes, et al., Conditional random fields as recurrent neural networks, in: IEEE International Conference on Computer Vision, IEEE, 2016, pp. 1529–1537.

[12] M. Zhao, Y. Wei, Y. Lu, K.K.L. Wong, A novel U-Net approach to segment the cardiac chamber in magnetic resonance images with ghost artifacts, Comput. Methods Programs Biomed. (2020) 105623.

[13] X. Zhu, Y. Wei, Y. Lu, M. Zhao, K. Yang, S. Wu, H. Zhang, K.K.L. Wong, Comparative analysis of active contour and convolutional neural network in rapid left-ventricle volume quantification using echocardiographic imaging, Comput. Methods Programs Biomed. (2020) 105914.

[14] K.K.L. Wong, A. Zhang, K. Yang, GCW-UNet segmentation of cardiac magnetic resonance images for evaluation of left atrial enlargement, Comput. Methods Programs Biomed. (2022) 106915.

[15] C. Liu, B. Chen, W. Shao, W.J. Zhang, Y. Zhang, K.K.L. Wong. Unraveling Attacks in Machine Learning-based IoT Ecosystems: A Survey and the Open Libraries Behind Them, Weblink, 2024 https://arxiv.org/abs/2401.11723.

[16] D.R. Katti, K.S. Katti, Cancer cell mechanics with altered cytoskeletal behavior and substrate effects: A 3D finite element modeling study, J. Mech. Behav. Biomed. Mater. 76 (2017) 125–134.

[17] P. Gut, M. Krieger, T. Lomax, D.C. Weber, J. Hrbacek, Combining rescanning and gating for a time-efficient treatment of mobile tumors using pencil beam scanning proton therapy, Radiother. Oncol. 160 (2021) 82–89. ISSN 0167-8140.

[18] C. Zhao, F.C. Lui, S. Du, D. Wang, Y. Shao, An earth mover's distance based multivariate generalized likelihood ratio control chart for effective monitoring of 3D point cloud surface, Comput. Ind. Eng. 175 (2023) 108911.

[19] V. Kumar, S. Patel, V. Baburaj, A. Vardhan, P.K. Singh, R. Vaishya, Current understanding on artificial intelligence and machine learning in orthopaedics – A scoping review, J. Orthopaed. 34 (2022) 201–206. ISSN 0972-978X.

[20] K.K.L. Wong, Z. Sun, J.Y. Tu, S.G. Worthley, J. Mazumdar, D. Abbott, Medical image diagnostics based on computer-aided flow analysis using magnetic resonance images, Comput. Med. Imaging Graph. 36 (7) (2012) 527–541.

[21] K. Ramesh, G.K. Kumar, K. Swapna, D. Datta, S.S. Rajest, A review of medical image segmentation algorithms, EAI Endorsed Trans. Perv. Health Tech. 7 (27) (2021) [Internet].

[22] K.K.L. Wong, Cybernetical Intelligence: Engineering Cybernetics with Machine Intelligence, John Wiley & Sons Limited, England, U.K., 2023. ISBN: 9781394217489.

[23] Y. Zhou, S. Du, M. Liu, X. Shen, Machine-fixture-pallet resources constrained flexible job shop scheduling considering loading and unloading times under pallet automation system, Journal of Manufacturing Systems 73 (2024) 143–158.

[24] Molin Liu, Jun Lv, Du. Shichang, Yafei Deng, Xiaoxiao Shen, Yulu Zhou, Multi-resource constrained flexible job shop scheduling problem with fixture-pallet combinatorial optimisation, Computers & Industrial Engineering 188 (2024) 109903.