Multi-Scale Tumor Localization Based on Priori Guidance-Based Segmentation Method for Osteosarcoma MRI Images

Baolong Lv 1, Feng Liu 1,*, Fangfang Gou 2 and Jia Wu 2,3, *

1 School of Information Engineering, Shandong Youth University of Political Science, Jinan 250102, China; lvbaolong2010@sina.com
2 School of Computer Science and Engineering, Central South University, Changsha 410017, China; gff8221@csu.edu.cn
3 Research Center for Artificial Intelligence, Monash University, Melbourne, VIC 3800, Australia
* Correspondence: liusdyu@sina.com (F.L.); jiawu0510@csu.edu.cn (J.W.)

Abstract: Osteosarcoma is a malignant osteosarcoma that is extremely harmful to human health. Magnetic resonance imaging (MRI) technology is one of the commonly used methods for the imaging examination of osteosarcoma. Due to the large amount of osteosarcoma MRI image data and the complexity of detection, manual identification of osteosarcoma in MRI images is a time-consuming and labor-intensive task for doctors, and it is highly subjective, which can easily lead to missed and misdiagnosed problems. AI medical image-assisted diagnosis alleviates this problem. However, the brightness of MRI images and the multi-scale of osteosarcoma make existing studies still face great challenges in the identification of tumor boundaries. Based on this, this study proposed a prior guidance-based assisted segmentation method for MRI images of osteosarcoma, which is based on the few-shot technique for tumor segmentation and fine fitting. It not only solves the problem of multi-scale tumor localization, but also greatly improves the recognition accuracy of tumor boundaries. First, we preprocessed the MRI images using prior generation and normalization algorithms to reduce model performance degradation caused by irrelevant regions and high-level features. Then, we used a prior-guided feature abdominal muscle network to perform small-sample segmentation of tumors of different sizes based on features in the processed MRI images. Finally, using more than 80,000 MRI images from the Second Xiangya Hospital for experiments, the DOU value of the method proposed in this paper reached 0.945, which is at least 4.3% higher than other models in the experiment. We showed that our method specifically has higher prediction accuracy and lower resource consumption.

Keywords: osteosarcoma; MRI image segmentation; multiscale tumor localization; prior guidance; AI-assisted diagnosis

MSC: 68T01

1. Introduction

Osteosarcoma is a primary malignant tumor of limb bone invasion that occurs in children and adolescents and has a very poor prognosis [1]. In developing countries, osteosarcoma has an average incidence of 0.0003% and is the most common malignant bone tumor other than multiple myeloma, accounting for 0.2% of malignancies. Due to its high degree of malignancy, distant metastasis can occur at an early stage, and at present, 20% of patients have pulmonary metastasis when diagnosed [2]. Lung metastasis is the main cause of treatment failure and death in patients with osteosarcoma, which can prolong the life cycle of osteosarcoma in the human body and further worsen the patient’s condition. Therefore, early diagnosis of osteosarcoma is very important for treatment and prognosis. In image omics, due to the magnetic resonance imaging (MRI) for osteosarcoma infiltration
range, the new judgments tend to be more precise to the size of the lesions, closer to the lesion resection after gross measurement results, and basic to the human body in the process of detection without damage [3]. Therefore, MRI images often become an important basis for doctors to diagnose osteosarcoma in imaging omics.

Clinical diagnosis of osteosarcoma in developing countries still faces many challenges. In most developing countries, the level of medical care is low and the ratio of doctors to patients is unbalanced, making it difficult to provide one-on-one and specialized medical services for patients. The diagnosis and treatment of osteosarcoma take a long time and cost a lot of money. For many poor families, they may be forced to give up treatment because of the high cost [4]. The lack of magnetic resonance imaging and other medical equipment and professional talents makes it very difficult to diagnose osteosarcoma in an early stage in a timely manner [5]. In addition, MRI images of osteosarcoma have a large amount of data. Each patient will generate over 500 MRI images at each diagnosis, and generally, only less than 20 images are valid. There is a lack of professional technical methods for automatic detection of these images, and relying only on doctors for manual examination will bring doctors too much redundant work [6–8]. In addition, the diagnosis of osteosarcoma requires a high degree of physician attention [9]. The long-term work of imaging doctors with a high load and a large number of redundant data inevitably leads to visual fatigue, missed diagnosis, misdiagnosis, and other problems [10–14].

The application of artificial intelligence to medical image processing has become a big boost. For example, the collection of image data provides a large number of samples [15,16]. The shape, location, and structure of the tumor region in osteosarcoma images vary greatly, in which the brightness model between different images is less interpretable for the image [17]. Therefore, the effect of the prior art in medical image segmentation has not yet reached people’s expectations.

The segmentation results obtained by the existing methods are quite different from the actual area of the tumor, and there is a problem of insufficient model generalization ability, which does not meet the expected standard for practical application in auxiliary medicine [18–20]. Knowledge-based segmentation has been applied to medical image processing by many researchers [21,22]. Due to the complexity of human anatomy and the systematic function of human body structure, such methods have achieved automatic segmentation of organs and tissues, as well as localization and identification of tumor regions [23]. However, manual intervention is still an anatomically necessary workflow. In addition, when processing MRI images of osteosarcoma, such methods do not solve the problem of multi-scale tumors well [24].

Based on this analysis, an osteosarcoma-assisted feature-enrichment network based on prior guidance is proposed for small shot segmentation (PESNet). First, the prior segmentation framework uses high-level features in the image to predict final semantic cues. It avoids the severe degradation of model performance caused by directly using the MRI original image features, and at the same time removes the irrelevant tumor regions in the MRI image and reduces the noise in the irrelevant regions. Then the isolated highlights are deleted and the algorithm normalized to process the pre-segmented regions generated prior. The obtained data set can speed up the training of the model. In addition, we used a priority-guided characteristic abdominal muscle network to perform low-shot segmentation of tumors of different sizes in the processed MRI images. Finally, we saved the segmented results in the form of graphics.

All contributions of this study are as follows:

1. The raw MRI images were processed using a priori generative algorithm. The prior generation can learn the key features of the image and make full use of the high-level features to provide clues for the final prediction. It avoids the severe degradation of model performance caused by directly using MRI raw image features.

2. We used preprocessing such as image normalization and removal of isolated bright spots to eliminate the interference of invalid areas and reduce resource waste.
(3) This paper proposed a prior-guided-based MRI image segmentation method for osteosarcoma (PESNet), which adds a priori generation and feature enrichment network to effectively improve the localization accuracy and segmentation accuracy of multi-scale tumors.

(4) The datasets used in this experiment were all from more than 200 real samples from the Second Xiangya Hospital. The results showed that the proposed segmentation method outperforms other methods. The prediction results of the model can be used as an auxiliary basis for doctors’ clinical diagnoses and improve the accuracy of diagnosis.

The other parts of this paper are as follows: The second chapter will introduce the existing research related to the auxiliary diagnosis of osteosarcoma. The third chapter introduces the design and implementation of the segmentation method. The fourth chapter is the analysis and discussion of the experiment. In Section 5, we present a conclusion and outlook.

2. Related Works

There are a number of artificial intelligence technologies used in medical imaging and decision-making systems to help doctors diagnose diseases [25–27]. For example, the Adam–Bashfort–Moulton prediction method [28] and the Hermite wavelet collocation method (HWM) [29] are used to describe the coronavirus with mathematical models; the epidemiological Susceptible-Infected-Removed model [30]; and the mathematical model for tumor invasion analysis [31]. We briefly introduce the existing osteosarcoma auxiliary diagnostic models.

Cases of osteosarcoma can be diagnosed by a combination of clinical, radiological, and histopathological examinations. Due to the complex structure of osteosarcoma, imaging plays an important role in the study of pathology. To this year, both in research and development at home and abroad, some medical image analysis software, using computer and image processing techniques such as image segmentation, 3-D visualization of medical data, and qualitative and quantitative analysis [32–34] has been used to assist doctors in providing accurate and comprehensive medical diagnosis, but the related studies mainly focus on the brain, heart, blood vessels, and bones, etc. However, there is limited medical application software for numerical-aided analysis of osteosarcoma.

At present, some scientific research institutions have developed related bone tumor medical analysis systems. The Department of Biomedical Informatics at Ohio State University has created a digital pathology image analysis system [35]. The system provides a data fusion function, which can connect the hospital information system and the image archive system, and can automatically analyze the pathological characteristics of osteosarcoma, neurocytoma, and other tumors. The Department of Mechanical Engineering of IIT developed the CAOS system named OrthoSYS [36]. The OrthoSYS system uses geometric theory to visualize the size and shape of tumors, and automatically identifies and annotates the anatomy of the 3D model bone. Moreover, OrthoSYS can automatically select the optimal model for prosthesis replacement at the bone tumor site. It provides a reference for doctors to choose the correct postoperative reconstruction method.

Linear filtering and nonlinear filtering are commonly used in MRI image filtering. Lee first proposed a filtering method based on the diffusion equation [37]. Lee filter believed that the multiplicative noise model could be approximated as a linear model, and set the same diffusion coefficients in all directions of the image, so the details of the image could not be retained well. Perona and Malik proposed an anisotropic diffusion equation [38] based on a partial differential equation, namely the p-M diffusion equation, which is an adaptive nonlinear filtering method that can enhance edge features.
However, osteosarcoma has extremely complex biological heterogeneity and individual differences in human tissue and organ structure, which makes the edge shape and texture characteristics of medical image of osteosarcoma very complicated. The research in this aspect has always been the focus and difficulty in medical image processing research. Therefore, image segmentation of each tissue structure in osteosarcoma image is of great significance, which is also the basis of visualization of osteosarcoma and quantitative analysis of clinical indicators. Chopp first proposed the narrowband level set method [39]. The basic idea of this method is to define a certain region whose distance from zero level set point is \( \varepsilon \) as a narrow band region, and update only the velocity and level set function values in this narrow band region during the curve evolution. This method not only reduces the computational complexity, but also improves the speed of segmentation. Ding Shaowei et al. [40] proposed a level set image segmentation method based on the combination of the fast marching method and narrowband method to segment tumor tissues in osteosarcoma sequence images. The algorithm can effectively improve the computing speed and segment the osteosarcoma tumor tissue in the image more accurately. However, there exists ambiguity of the edge in the osteosarcoma image and a problem of edge leakage in the evolution process of the level set method.

Nasor et al. [41] proposed a segmentation method for osteosarcoma combining k-means and other image processing techniques. This method can segment osteosarcoma without considering the changes in intensity, texture, and location. It was tested in combination with 50 image data and successfully segmented tumor regions in MRI images. However, the algorithm is complex, computationally intensive, and may destroy the region boundaries in image segmentation. David Joon Ho et al. [42] described deep interactive learning (DiLaL), and after 7 h of DiLaL labeling, the necrosis rate within the range of expected interobserver variation rate could be successfully estimated. In addition, Huang et al. [43] proposed an automatic segmentation method for MRI images of osteosarcoma.

However, edge features are difficult to fully utilize [44–46]. In general, the accuracy and precision of segmentation need to be further improved, and there is still a large error segmentation rate. We proposed an automatic segmentation method for MRI images of osteosarcoma, a feature enrichment network based on the prior guided sparse shot segmentation method. This method improves the overall accuracy of image segmentation edge optimization and saves computing resources required by hardware devices.

3. Methods

With the development of artificial intelligence in the medical field, image processing technology plays an increasingly important role in the examination, treatment, and prognostic management of osteosarcoma [47–49]. Our method has wider applicability and lower hardware and software requirements. As shown in Figure 1, this paper proposes a prior guidance-based segmentation method for osteosarcoma MRI images. It is mainly divided into three parts. The first is to generate a priori information for MRI images. The resulting images are then preprocessed. That is, after removing isolated bright spots and normalizing, a new dataset is generated for model training. Finally, the detailed description of the segmentation method is given. During model training, the input dataset is first collected by a feature enrichment module and then segmented by a network consisting of convolutional blocks and classification heads. Finally, the segmentation result of the MRI image is obtained and stored in the form of an image.

This paper is divided into two Sections. Section 3.1 describes the few-shot cutting task and the prior-guided feature enrichment network used for the MRI image cutting of osteosarcoma, and Section 3.2 describes the specific segmentation process of MRI images and the graphic results obtained.
Figure 1. A holistic architecture for segmentation of osteosarcoma MRI images.

The main symbols in this chapter are shown in Table 1.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Paraphrase</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_i$</td>
<td>The $i$-th support set</td>
</tr>
<tr>
<td>$q_i$</td>
<td>The $i$-th query set</td>
</tr>
<tr>
<td>$S, Q$</td>
<td>Origin data set</td>
</tr>
<tr>
<td>$I_Q$</td>
<td>The query image</td>
</tr>
<tr>
<td>$C_n$</td>
<td>The unknown class</td>
</tr>
<tr>
<td>$C_t$</td>
<td>The test class</td>
</tr>
<tr>
<td>$A = {I_Q, M_Q}$</td>
<td>Query sample set</td>
</tr>
<tr>
<td>$M_Q$</td>
<td>Query the Mask of the set</td>
</tr>
<tr>
<td>$X_Q, X_S$</td>
<td>Original query feature</td>
</tr>
<tr>
<td>$B = [B_1, B_2, ..., B_n]$</td>
<td>Average $n$ different space sizes of the pool</td>
</tr>
<tr>
<td>$X_{Q, FEM}$</td>
<td>Subquery feature</td>
</tr>
<tr>
<td>$L_1$</td>
<td>Loss value</td>
</tr>
</tbody>
</table>

3.1. Few-Shot

The Few-Shot method is divided into $S$ support set and $Q$ query set. The core idea is to segment the unknown $C_n$ from each query MRI image $I_Q$ in the MRI image collection of osteosarcoma given $K$ samples of the $S$ set.

First, we used a prior feature-based network, as shown in Figure 2, which is also known as a prior-guided feature enrichment network. CNN model extracts osteosarcoma features from MRI images by supporting and querying image sharing. The number of
channels is reduced to 256 by 1 \times 1 convolution blocks to handle the middle-level support and query features.

\[
X_{1} = f(I_{1}), \quad X_{2} = f(I_{2}) \land M.
\] (1)

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{The preprocessing process of MRI Image.}
\end{figure}

The query features are then enriched using convolutional blocks. Finally, there is a classification module, which mainly consists of the softmax function.

3.2. Image Segmentation of Osteosarcoma

The segmentation of osteosarcoma can be divided into data preprocessing and image analysis.

To further improve the detection accuracy, three strategies were established:

(1) Prior Generation. In contrast to the adverse effects of high-level features on the performance of few-shot segmentation, prior segmentation frameworks use these features to provide semantic cues for the final prediction. We performed prior generation processing on MRI images of osteosarcoma to reduce the interference of invalid active segmentation regions on the final prediction, thereby improving the efficiency of image processing.

(2) Pretreatment. We further preprocessed the MRI image generated after prior generation, and processed the mask and prior generation results respectively by deleting isolated highlights and the normalization algorithm to speed up model training and save computing resources.

(3) Image analysis and segmentation. The segmentation model in this paper is a feature enrichment network based on prior guidance. When training the model, the MRI image of osteosarcoma and its preprocessed mask were input into the network to confirm loss function, and the error segmentation rate of the osteosarcoma image was reduced through repeated training.

3.2.1. Prior Generation

In MRI images of osteosarcoma, the background region occupies most of the space. However, this information is useless for both model training and image segmentation. They result in inefficient use of resources such as memory. In addition, since the gray values of the tumor areas are similar, it is also possible to interfere with the final result. Therefore, we chose to clear this type of area.

Experiments on CANet [50] showed that high levels of features in MRI images lead to performance degradation. The performance of middle-level features is better, and the specificity of high-level features contained in semantic information is stronger than that of middle-level features, indicating that the former has a more negative effect on the generalization ability of invisible classes. However, advanced features can directly provide the semantic information of the training class and make a greater contribution to the recognition of pixels belonging to the class. Compared with medium-level features, training loss is reduced. Thus, this behavior leads to a preference for process pairs predicted by the model. Lack of generalization ability is not conducive to accurate segmentation of actual tumor areas.
By converting ImageNet’s pre-trained advanced features containing semantic information into a prior image, the mask is the pixel the probability of belonging to the target. The calculation formula is as follows:

$$X_Q = f(I_Q), \ X_S = f(I_S) \odot M.$$  \hspace{1cm} (1)

Here $\odot$ is Hadamard product, $X_Q$ and $X_S$ represent advanced query and supported features, $f$ represent trunk network, $M$ represents binary supported mask, and $I_Q$, $I_S$ represent the input query and supporting MRI images, respectively. Please note that the ReLU function was used as the activation function to output the medium $f$.

Specifically, this study defined $X_Q$ and $Y_Q$ as the corresponding relational masks $X_Q$ and $X_S$, respectively. The larger the value of $Y_Q$, the higher the corresponding relationship. That is, it queries the greater probability in the target region of the MRI image [51], so we need to set the background of irrelevant tumor feature information in the support feature to zero. When generating $Y_Q$, we need to calculate the pixel-level cosine similarity $\cos(x_q, x_s) \in \mathbb{R}$ between $x_q \in X_Q$ and $x_s \in X_S$ feature vectors, and the calculation formula is as follows:

$$\cos(x_q, x_s) = \frac{x_q^T x_s}{||x_q|| \cdot ||x_s||}, \ q, s \in \{1, 2, \ldots, hw\}. \hspace{1cm} (2)$$

For each $x_q \in X_Q$, the maximum similarity between support pixels can be represented by the corresponding $c_q \in \mathbb{R}$. The specific description is as follows:

$$c_q = \max_{s \in \{1, 2, \ldots, hw\}} \cos(x_q, x_s), \hspace{1cm} (3)$$

$$C_Q = [c_1, c_2, \ldots, c_{hw}] \in \mathbb{R}^{hw \times 1}. \hspace{1cm} (4)$$

After obtaining $c_q$ and $C_Q$ through calculation, we generate prior query feature $Y_Q$ by remodeling $C_Q \in \mathbb{R}^{hw \times 1}$ into $Y_Q \in \mathbb{R}^{h \times w \times 1}$, i.e., the pre-segmented image.

The key to the prior generation method is to use fixed, advanced features and use the maximum value in the similarity matrix as the prior query feature; compared with using the original image directly, the prior generation method can clearly understand the interference of noise in irrelevant regions on segmentation results. The model pays more attention to the actual tumor area when predicting osteosarcoma MRI images, so it has a better refinement effect on the tumor boundary.

### 3.2.2. Data Preprocessing

We further altered the prior generation of osteosarcoma MRI image preprocessing, highlights, and the normalization algorithm by deleting isolated highlights, respectively, in MRI images of the mask and prior to generating results for processing, thus eliminating singular sample data leading to negative effects, and speeding up the optimal solution of the gradient descent, improving the precision of which saves computing resources at the same time. Figure 2 illustrates the prior generation and data preprocessing process for MRI images. It includes several steps of the prior generation, clear outliers, the establishment of a trust zone, and normalization, as follows:

1. **Delete isolated highlights**

   After prior generation of MRI images, the real tumor region cannot be established, and there are usually some isolated bright spots in the lesion region, which increases the difficulty of image segmentation to predict accurate results, so we chose to delete these isolated bright spots.

2. **Determine the trusted region**

   The osteosarcoma dataset after the deletion of isolated bright spots has different gray values, lesion areas, and shapes, but lesion areas are often connected. To satisfy the premise that the pre-segmented area must contain the lesion area, in the horizontal direction, the
maximum value of the coordinates of the bright spot was set to \( \max(Y_Q) \), and the minimum value was set to \( \min(Y_Q) \). Similarly, in the vertical direction, the maximum and minimum values were calculated by \( \max(Y_Q) \) and \( \min(Y_Q) \) respectively. Finally, we calculated the required area as: \( [\max(X_Q): \min(X_Q), \max(Y_Q): \min(Y_Q)] \).

(3) Normalization

We normalized \( Y_Q \) to between 0 and 1 using min-max normalization \(^52\). \( \epsilon \) was set to \( 1 \times 10^{-7} \) in our experiment.

\[
X_Q = \frac{X_Q - \min(X_Q)}{\max(X_Q) - \min(X_Q)} + \epsilon, \quad (5)
\]

\[
Y_Q = \frac{Y_Q - \min(Y_Q)}{\max(Y_Q) - \min(Y_Q)} + \epsilon. \quad (6)
\]

The key to the prior generation method of our experiment is to use the advanced features fixed in the osteosarcoma graph to obtain the prior mask by the similarity matrix. This method is quite simple and effective.

3.2.3. Image Analysis and Segmentation

The segmentation of the tumor network model of osteosarcoma is an a priori guided feature enrichment network (PESNet), as shown in Figure 3. It mainly includes an a priori generation module and a feature enrichment module (FPE). First, the query set, support set, and support mask were input into the Prior generation module. The query set has more prior features to generate the corresponding prior query set. Then, the feature information of the globally pooled support set and support mask is input into the feature enrichment module together with the Prior Query Feature. The FPE module again decomposes and constructs the multi-scale MRI images of osteosarcoma. Among them, at each scale, query features, support features, and prior masks interact. Furthermore, PESNet exploits hierarchical relationships to enrich coarse-grained features by extracting essential information from fine-grained features through self-directed information paths. Through horizontal and vertical optimization, the model obtains features of different scales as new query features.

![Figure 3. PESNet segmentation model.](image-url)
Query and support features. The priority mask is the input to this module. The refined query features from the support features are the output of this module [41]. Proceed as follows:

1. Inter-source enrichment: It mainly maps osteosarcoma MRI images to different scales, and then interacts with the query, support features, and prior masks of the model.
2. Inter-scale interaction: It mainly transfers information between some features of different scales.
3. Information concentration: It combines features of different scales, eventually generating refined query features, providing a basis for determining the final region and location of the tumor query features.

We are sorted in descending order $B^1 > B^2 > ... > B^n$. After the query features $X_Q$ were input, $n$ sub-query features $X_Q^{EM} = [X_Q^1, X_Q^2, ..., X_Q^n]$ of different spatial sizes were generated from the features of the MRI image query set of osteosarcoma using adaptive averaging pool processing. The n space size enables the global mean set support feature $X_S \in \mathbb{R}^{1 \times 1 \times c}$ to be extended to n feature maps $X_Q^{EM} = [X_Q^1, X_Q^2, ..., X_Q^n] (X_S^i \in \mathbb{R}^{B^i \times B^i \times c})$, and prior $Y_Q \in \mathbb{R}^{B^i \times w \times h}$ is adjusted accordingly to $Y_Q^{EM} = [Y_Q^1, Y_Q^2, ..., Y_Q^n] (Y_Q^i \in \mathbb{R}^{B^i \times B^i \times c})$.

Then, for $i \in \{1, 2, ..., n\}$, we connected $X_i Q$, $X_i S$, and $Y_i Q$ and processed each connected feature with convolution to generate the combined query feature $X_{Q,m}^i \in \mathbb{R}^{B^i \times B^i \times c}$ of osteosarcoma:

$$X_{Q,m}^i = F_{1 \times 1} (X_Q^i \oplus X_S^i \oplus Y_Q^i).$$  \hspace{1cm} (7)

Among them, the combined $1 \times 1$ convolution is represented by $F_{1 \times 1}$. It has $c = 256$ output channel features.

Where $F_{1 \times 1}$ represents the convolution of $c = 256$ output channel features with a size of $1 \times 1$ after merging.

Interactions between scales. It is worth noting that blurred segmentation boundaries have always been a big problem in image segmentation tasks. The self-adaptive transmission of information from fine features to rough paths helps to establish hierarchical relationships in our feature-rich module and refine and differentiate tumor region boundaries.

M in the circle of Figure 3 represents the merging module between scales. This process can be written as

$$X_{Q,m}^{\text{new}} = M (X_{Q,m}^{\text{Main},i}, X_{Q,m}^{\text{Aux},i}).$$  \hspace{1cm} (8)

Here, $X_{Q,m}^{\text{Main},i}$ is the main feature and $X_{Q,m}^{\text{Aux},i}$ is the auxiliary feature of tumor query of the i-th scale $B^i$. For example, in a top-down finite element model with interscale interaction pairs, the finer feature (auxiliary) $X_{Q,m}^{i-1}$ is required to provide additional information for the coarse feature (main) $X_{Q,m}^i (B^{i-1} > B^i, i \geq 2)$. In this case, $X_{Q,m}^{\text{Aux},i} = X_{Q,m}^{i-1}$ and $X_{Q,m}^{\text{Main},i} = X_{Q,m}^i$. Other alternatives to interscale interaction include bottom-up pathways that use information on auxiliary features to enrich more fine-grained tumor main features, and dual variants.

Interscale merging. The specific structure of module M is shown in the upper right corner of Figure 3.

Information concentration. After the interaction between scales, n refined feature graphs $X_{Q,new}^i, i \in \{1, 2, ..., n\}$ are obtained. Finally, we generate the output query feature $X_{Q,new}^i \in \mathbb{R}^{B^i \times w \times h}$.

$$X_{Q,new}^i = F_{1 \times 1} (X_{Q,new}^1 \oplus X_{Q,new}^2 \oplus Y_{Q,new}^n).$$  \hspace{1cm} (9)

In conclusion, by combining the combined supporting features with the priority mask to query features of different spatial sizes, the model learned to enrich the query features adaptively with the effective information from the supporting features of MRI images of osteosarcoma under the guidance of the priority mask. In addition, the main features were supplemented by the conditional information provided by the auxiliary features of vertical
interscale interactions. Thus, the finite element method yields greater performance gains at baseline than other feature-enhanced designs.

According to the feature that the lesion area in the MRI image is small relative to the overall image, the loss function of this study selects the cross-entropy loss [53]. The loss function proposed in this paper is as follows:

\[ L = \frac{\sigma}{n} L_1 + L_2. \]  

(10)

Among them, it is used to balance the effect of intermediate supervision. After the experiment, we set the weight parameter \( \sigma = 1.0 \).

In this paper, MRI images of osteosarcoma in different parts of the human body were used to train the model, and three types of image segmentation results of sagittal plane, coronal plane, and cross section were obtained. In the clinical diagnosis of osteosarcoma, early and timely detection and location of osteosarcoma is the key to the successful treatment of patients [54]. Our segmentation method can not only accurately detect the location of osteosarcoma, but also requires less medical hardware facilities, which can reduce the cost of medical treatment. The physician uses the segmentation results and the results of the final focal results as an auxiliary basis for the diagnosis of osteosarcoma, which helps to provide accurate diagnosis of osteosarcoma.

4. Experiments and Results

All data in this paper came from the Ministry of Mobile Health Education-China Mobile United Medical Laboratory and the Second Xiangya Hospital of Central South University [55]. Among them, more than 80,000 MRI images of more than 200 patients in the hospital in recent years were included. In the experiments in this paper, we divided the training set and test set in a ratio of 7:3.

4.1. The Evaluation Index

For the evaluation of the performance of each model, our metrics include the following: Accuracy (Acc), Precision, Recall, F1-score, IOU, HM, and DSC. A confounding matrix consisting of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) was used to explain the performance of the network [56].

Acc was defined as follows:

\[ Acc = \frac{TP + TN}{TP + TN + FP + FN}. \]  

(11)

Precision (Pre) refers to the proportion of positive samples, that is, the ratio of the number of correctly retrieved samples to the total number of retrieved samples. It was defined as:

\[ Pre = \frac{TP}{TP + FP}. \]  

(12)

For the segmentation results of osteosarcoma, we tried to avoid missed diagnosis by maximizing the recall rate [57]. Recall (Re) was defined as follows:

\[ Re = \frac{TP}{TP + FN}. \]  

(13)

We defined F-score as follows:

\[ F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}. \]  

(14)
DSC and IOU are the ratios of similarity coefficient and overlapping area, respectively. Their values were all \([0, 1]\). We defined them as follows:

\[
DSC = \frac{2 \times |S_1 \cap S_2|}{|S_1| + |S_2|},
\]

\[
IOU = \frac{|S_1 \cap S_2|}{|S_1 \cup S_2|}.
\]

### 4.2. Training Strategy

We used MSFCN [58], MSRN [59], FCN [60], FPN [61], Unet [62], and PESNet for comparative experiment. Before the training segmentation model, to avoid excessively focus on insignificant features, we expanded the data set by enlarging (narrowing) images, rotation, and flip.

The training division neural network was trained for 300 epochs. We set Adam as the optimizer. In addition, the initial learning rate was set to 0.0055. After each round of learning rate, the change in gradient updates was more adaptable.

### 4.3. Segmentation Effect Evaluation

In our method, we found that the active regions containing bone tumors were only part of the image, while the other regions contained no active information, and it was possible to interfere with the resulting results. Therefore, we chose to use the prior generation algorithm, which classifies pixels based on the grayscale value of the pixels and removes this type of area. As shown in Figure 4, the left image is an MRI of the bone tumor, which clearly shows the boundary between different tissues. The right image is the MRI image after prior processing. It can be seen from the figure that the interference of invalid information on the cutting of the tumor region is removed after processing by the prior generation algorithm. Therefore, the segmentation model is able to focus more on the cutting of the actual tumor region.

![Figure 4. Comparison of data sets before and after processing.](image)

Figure 5 is a comparison of the effect of model segmentation before and after dataset processing. The left column is the true label value, the middle column is the result of the bone tumor MRI image without prior generation and processing, and the rightmost column is the prediction map after data set processing. Before optimization, the segmentation results of MRI images were incomplete, and there were cases where tumor location was misidentified. The optimized semantic segmentation results are closer to ground truth labels. As can be seen from the figure, after the prior generation processing, the segmentation effect of the PESNet model was significantly improved.
Figure 5. The comparison of image segmentation effect between models before and after prior generation.

Figure 6 shows the segmentation effect of different models on MRI images of bone tumors. The first is the tumor area of the tumor MRI, which is a real lesion area. The middle five is the image segmentation result of the control model, and the final column is the final result of the segmentation of the image for our PESNet method. We finally found that the two groups of PESNet were able to better divide the bone tumor, and the result was the most fitting of the actual lesion area. The result of segmentation is not only in shape but in the actual tumor area, but also with the location of the actual tumor, indicating that our model is better than other methods in terms of accuracy and ability of the right. The segmentation results can provide a more powerful auxiliary foundation for the diagnosis of the doctor, while improving the correctness of the diagnosis.

<table>
<thead>
<tr>
<th>Model</th>
<th>Acc</th>
<th>Pre</th>
<th>Re</th>
<th>F1</th>
<th>IOU</th>
<th>HM</th>
<th>DSC</th>
<th>Params</th>
<th>FLOPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSFCN</td>
<td>0.992</td>
<td>0.881</td>
<td>0.936</td>
<td>0.906</td>
<td>0.874</td>
<td>0.170</td>
<td>0.906</td>
<td>20.38M</td>
<td>1524.3G</td>
</tr>
<tr>
<td>MSRN</td>
<td>0.988</td>
<td>0.839</td>
<td>0.902</td>
<td>0.866</td>
<td>0.887</td>
<td>0.229</td>
<td>0.866</td>
<td>14.27M</td>
<td>1461.2G</td>
</tr>
<tr>
<td>FCN-8s</td>
<td>0.974</td>
<td>0.941</td>
<td>0.873</td>
<td>0.901</td>
<td>0.772</td>
<td>0.203</td>
<td>0.876</td>
<td>134.3M</td>
<td>190.08G</td>
</tr>
<tr>
<td>FCN-16s</td>
<td>0.990</td>
<td>0.922</td>
<td>0.882</td>
<td>0.900</td>
<td>0.824</td>
<td>0.326</td>
<td>0.859</td>
<td>134.3M</td>
<td>190.55G</td>
</tr>
<tr>
<td>FPN</td>
<td>0.989</td>
<td>0.919</td>
<td>0.924</td>
<td>0.921</td>
<td>0.852</td>
<td>0.186</td>
<td>0.883</td>
<td>88.63M</td>
<td>141.45G</td>
</tr>
<tr>
<td>UNet</td>
<td>0.991</td>
<td>0.918</td>
<td>0.929</td>
<td>0.924</td>
<td>0.867</td>
<td>0.100</td>
<td>0.892</td>
<td>17.26M</td>
<td>160.16G</td>
</tr>
<tr>
<td>Ours</td>
<td>0.995</td>
<td>0.940</td>
<td>0.945</td>
<td>0.945</td>
<td>0.898</td>
<td>0.102</td>
<td>0.945</td>
<td>10.82M</td>
<td>369.38G</td>
</tr>
</tbody>
</table>

Figure 6. Segmentation effect of each model on MRI images of osteosarcoma.
To more clearly evaluate the performance of different methods, we quantified the results predicted by the model. We compared the performance on the osteosarcoma dataset as shown in Table 2. As can be seen from the data in the table, the PESNet model showed good performance. Except for the accuracy and HM, the model was slightly behind second place. Accuracy, recall rate, F1, and IOU were increased by 0.4%, 0.9%, 2.1%, and 1.1%, respectively. The DSC value is very useful for the model, and our method increased by about 3.9% to 0.945. This shows that the model is able to segment more accurately, and the overall segmentation performance is significantly improved, thus providing a more valuable reference for medical decision-making.

Table 2. Comparison of performance evaluation indexes of MRI images of patients with osteosarcoma in different segmentation models.

<table>
<thead>
<tr>
<th>Model</th>
<th>Acc</th>
<th>Pre</th>
<th>Re</th>
<th>F1</th>
<th>IOU</th>
<th>HM</th>
<th>DSC</th>
<th>Params</th>
<th>FLOPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSFCN</td>
<td>0.992</td>
<td>0.881</td>
<td>0.936</td>
<td>0.906</td>
<td>0.874</td>
<td>0.170</td>
<td>0.906</td>
<td>20.38M</td>
<td>1524.3G</td>
</tr>
<tr>
<td>MSRN</td>
<td>0.988</td>
<td>0.839</td>
<td>0.902</td>
<td>0.866</td>
<td>0.887</td>
<td>0.229</td>
<td>0.866</td>
<td>14.27M</td>
<td>1461.2G</td>
</tr>
<tr>
<td>FCN-8s</td>
<td>0.974</td>
<td>0.941</td>
<td>0.873</td>
<td>0.901</td>
<td>0.772</td>
<td>0.203</td>
<td>0.876</td>
<td>134.3M</td>
<td>190.08G</td>
</tr>
<tr>
<td>FCN-16s</td>
<td>0.990</td>
<td>0.922</td>
<td>0.882</td>
<td>0.900</td>
<td>0.824</td>
<td>0.326</td>
<td>0.859</td>
<td>134.3M</td>
<td>190.55G</td>
</tr>
<tr>
<td>FPN</td>
<td>0.989</td>
<td>0.919</td>
<td>0.924</td>
<td>0.921</td>
<td>0.852</td>
<td>0.186</td>
<td>0.883</td>
<td>88.63M</td>
<td>141.45G</td>
</tr>
<tr>
<td>UNet</td>
<td>0.991</td>
<td>0.918</td>
<td>0.929</td>
<td>0.924</td>
<td>0.867</td>
<td>0.100</td>
<td>0.892</td>
<td>17.26M</td>
<td>160.16G</td>
</tr>
<tr>
<td>Ours</td>
<td>0.995</td>
<td>0.940</td>
<td>0.945</td>
<td>0.945</td>
<td>0.898</td>
<td>0.102</td>
<td>0.945</td>
<td>10.82M</td>
<td>369.38G</td>
</tr>
</tbody>
</table>

We trained 300 epochs in total, choosing 50 epochs for a comparative analysis of the epochs (an epoch was selected every six rounds), as shown in Figure 7. As we can see, with the increase of the epoch, the accuracy of each model was shaken, with MSRN the largest, and the change curve of our method being the most gentle. After 120 epochs, the accuracy of other models was stable in addition to MSRN in the middle of a certain epoch. At the highest number, PESNet (ours), the ultimate stability was 0.995. The accuracy ranking is PESNet (ours) > MSFCN > UNet > FPN > MSRN.

Figure 7. The accuracy variation of different models.

At the same time, we selected four models with our method in Figure 8; the F1 of MSRN and MSFCN models fluctuated greatly during the first 180 epochs of training. After that, all models reached a stable state. Finally, although the change curve of the F1 value proposed by our method is not the gentlest, it is also the highest among all models.
At the same time, we selected four models with our method in Figure 8; the F1 of MSRN and MSFCN models fluctuated greatly during the first 180 epochs of training. After that, all models reached a stable state. Finally, although the change curve of the F1 value proposed by our method is not the gentlest, it is also the highest among all models.

We selected several methods to compare DSC with our method. As shown in Figure 9, the fluctuation of our method was small. Although there was an obvious decline when the epoch was 120, the DSC value kept increasing with the increase of the epoch. Finally, its DSC value was always the highest. The ranking of DSC value was PESNet (ours) > MSFCN > UNET > FPN > MSRN. This indicates that the segmentation results obtained by our model have the highest sample similarity and best fit the real tumor region.

According to Figure 10, we can observe our proposed method was the best in DSC. In addition, our method retains fewer parameters, only 10.82 m, far less than 134.3 M in the FCN method. This means that our approach consumes fewer system resources during training, reducing the difficulty of training. The UNet model and MSFCN models are weaker than that our proposed method, but they also have better performance and fewer parameters than other models.
Finally, the comparison between FLOPs and DSC of different model methods is shown in Figure 11. The results showed that the relatively low VALUE of FLOPs in PESNet can significantly improve the accuracy and occupy less system memory resources, and, at the same time, it does not need to increase computing cost too much to achieve cost-effective time-space transaction, which reduces the overall requirements on hardware facilities. The DSC value of MSFCN is also very high, but this model requires a very large calculation cost, and the overall effect is not as good as our method.

![Figure 10. Comparison of parameters and DSC in different osteosarcoma models.](image)

**Figure 10.** Comparison of parameters and DSC in different osteosarcoma models.

5. Conclusions

This paper used more than 80,000 osteosarcoma MRI images as a dataset and proposed an osteosarcoma MRI image segmentation model (PESNet). The method is a feature-rich network based on prior guidance, including prior generation, data preprocessing, segmentation model, and result graphing. Experimental results showed that PESNet had a more significant performance in osteosarcoma MRI image segmentation. It reduced the segmentation error rate of traditional methods and saved computational resources to a certain extent.

Currently, this method can only be applied to the analysis of MRI 2D images. In the future, with the development of technology and the expansion of data sets, the identification...
and localization of osteosarcoma in 3D images, as well as the calculation of tumor area volume, will be the focus of research.

Author Contributions: Conceptualization, F.L. and J.W.; Data curation, B.L.; Formal analysis, F.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in the Shandong Social Science Plan Funds Projects, 21CZHJ16; Focus on Research and Development Projects in Shandong Province (Soft Science Project), 2021RKY02029.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: “Data Availability” statement data used to support the findings of this study are currently under embargo while the research findings are commercialized. Requests for data, 12 months after publication of this article, will be considered by the corresponding author. Availability of data and materials: all data analyzed during the current study are included in the submission.

Conflicts of Interest: The authors declare no conflict of interest.

References

4. Yu, G.; Chen, Z.; Wu, J.; Tan, Y. Medical decision support system for cancer treatment in precision medicine in developing countries. Peer J. Appl. 2021, 186, 115725. [CrossRef]


30. Gao, W.; Baskonus, H.M. Deeper investigation of modified epidemiological computer virus model containing the Caputo operator. *Chaos Solitons Fractals* 2022, 158, 112050. [CrossRef]


