Academic Journal of Computing & Information Science
ISSN 2616-5775 Vol. 8, Issue 10: 90-98, DOI: 10.25236/AJCIS.2025.081012

Liver Health: A Deep Learning-Based Auxiliary Diagnostic System for Liver Tumor Segmentation in CT Images

Wei Gao

School of Computing, Beijing Institute of Technology, Zhuhai, China, 519088

Abstract: *Efficient and accurate diagnosis from CT images is a crucial step in liver cancer diagnosis and treatment. Currently, liver cancer diagnosis mainly relies on manual work, which is heavily dependent on doctors' experience and susceptible to their subjective factors. This paper will explore in-depth automatic segmentation methods for the liver and liver tumors in CT images based on deep learning. By implementing architectures such as Unet and its variants, and simultaneously calling the Qwen-VL-Max multimodal large model to evaluate segmentation results, and using the PYQT framework to build an interface, the system achieves precise segmentation of liver and tumor regions in CT images. Thus, an efficient, easy-to-operate, and highly interactive liver tumor segmentation system is constructed to provide technical support for doctors' auxiliary diagnosis.*

Keywords: *Deep Learning, U-Net and Its Variants, Software System*

1. Introduction

With the development of social economy, people's living standards have undergone tremendous changes, especially in terms of diet and lifestyle, which have become increasingly rich and diverse. However, this has been accompanied by a gradual increase in the incidence of liver diseases, especially liver tumors. According to the latest GLOBOCAN data [1], which summarizes and analyzes the global statistics on cancer incidence and mortality, the data shows that in recent years, the number of new cases of 36 types of cancer worldwide has reached as high as 19.96 million, with approximately 9.74 million deaths. Among them, the incidence of liver cancer accounts for about 4.7%, ranking sixth globally, and the mortality rate accounts for about 7.8%, ranking fourth globally. It is a malignant tumor with extremely high incidence and mortality rates. The latest data shows that the annual number of new liver cancer cases in China ranks fourth among malignant tumors, and the number of deaths ranks second, resulting in a heavy disease burden. Therefore, it is crucial to detect liver tumors as early as possible and take corresponding treatment measures, which plays a vital role in improving patients' survival rates [2]. At present, the diagnosis and treatment of most liver diseases mainly rely on manual diagnosis. Doctors need to observe and segment abdominal CT images based on clinical experience, and then diagnose liver problems based on the results. This traditional analysis method is time-consuming, labor-intensive, and inefficient, making it difficult to ensure accuracy in large-scale patient data. At the same time, it is heavily dependent on doctors' experience and susceptible to subjective factors, which may easily lead to delayed treatment. It can no longer meet the growing diagnostic needs. Therefore, there is an urgent need for automated liver tumor segmentation auxiliary diagnostic software systems to be applied in the clinical practice of liver auxiliary diagnosis [3][4].

In recent years, the development of deep learning technology has brought revolutionary changes to the field of medical diagnosis. These technologies can use advanced algorithms to process massive amounts of data, thereby realizing the automated processing of complex diagnostic tasks, optimizing diagnostic processes, and reconstructing medical services [5][6]. At present, scholars at home and abroad have conducted a lot of research on computer-aided diagnosis technology. In the field of liver cancer diagnosis software systems, Li Shufeng [7] and others proposed a design idea and implementation framework of a medical image automatic diagnosis system integrating deep learning technology, which has certain practical significance in terms of sensitivity and specificity; Li Mingyang [8] proposed to develop an interface based on PyQt5, built a system for automatic segmentation of liver tumor CT images, which can load and display CT images, achieve accurate segmentation of liver tumors by embedding the FRTrans-Unet network model, and save the segmentation results, providing doctors with a convenient and efficient auxiliary diagnostic tool; Wang Rongmiao [9] proposed to build a platform for liver 3D visualization, programmed the researched algorithms using C++, and initially realized functions such as image sequence reading, browsing, liver parenchyma segmentation, intrahepatic vessel segmentation, and 3D reconstruction. The above research works can help doctors efficiently segment liver tumors and improve diagnostic efficiency. However, the designed systems have problems such as cumbersome operation, unintuitive interface design that makes it inconvenient for doctors to use, single functional modules (only capable of image segmentation without providing preliminary diagnosis to doctors), and inability to save and record results.

Nowadays, in the field of medical diagnosis, we have entered the stage of using artificial intelligence for auxiliary diagnosis, and computer systems for auxiliary diagnosis have received attention and relatively extensive applications in the medical field. Based on the problems existing in the above research works, this paper designs and develops a new liver cancer auxiliary diagnostic software - LiverHealth. The contributions of this paper are as follows:

(1) The software builds a user-friendly graphical interface through the QT framework, encapsulates the complex deep learning segmentation process into easy-to-operate tools, and designs an intuitive interface for doctors' actual use, enabling doctors to directly import CT images and operate by clicking the corresponding functional modules.

(2) After importing CT images, doctors can directly perform liver tumor segmentation using the software. In addition, a database is designed, allowing doctors to choose whether to save the results into the database. Innovatively, the software calls a large model to generate a preliminary diagnostic report based on the segmentation results, which can be exported as a PDF file and saved locally.

(3) The software uses a total of 4 segmentation models, namely Unet, Unet++, Unet3+, and Segformer. Their evaluation indicators on the LIST dataset used in this paper all reach 0.8 or above, indicating excellent performance of the models. These functional designs and the excellence of the models indicate that the software system designed in this paper can assist doctors in liver cancer auxiliary diagnosis.

2. Software Architecture

With the development of computed tomography technology, the number of CT images is constantly increasing. However, CT images have characteristics such as low contrast and similar grayscale values between organ tissues, which put great pressure on doctors when processing large medical image datasets. At present, computer-aided diagnosis technology has been applied in China to a certain extent. It has the characteristics of fast speed and high resolution, thus achieving wide application in the medical field, but its use cost is relatively high. Therefore, this chapter will introduce in detail the proposed deep learning-based liver tumor segmentation auxiliary diagnostic system. By developing a user-friendly graphical interface through the QT framework, complex deep learning segmentation processes are encapsulated into easy-to-operate tools, enabling doctors to conveniently use the system for liver and tumor segmentation and enhancing the practicality of the technology.

2.1 Overall System Design

The functions of this system include importing local liver CT slices, image processing and liver tumor segmentation, storing segmentation results in a MYSQL database, reading the database, generating diagnostic reports based on segmentation results, and exporting PDF files. Firstly, preprocessing is performed on the LIST dataset to generate a newdata file. Using this dataset, relevant onnx models are trained with GPU in the Baidu Wenxin Yiyan environment. Then, the onnx models are imported into the built model, and the software front-end is built using PYQT5. Finally, the front-end interface is run under the main function.

This section will detail how doctors use the system for liver tumor segmentation and auxiliary diagnosis. In this process, doctors first need to import CT images, and then press the corresponding buttons to trigger relevant processing operations as needed. Specifically, doctors can operate according to the process shown in Figure 1.

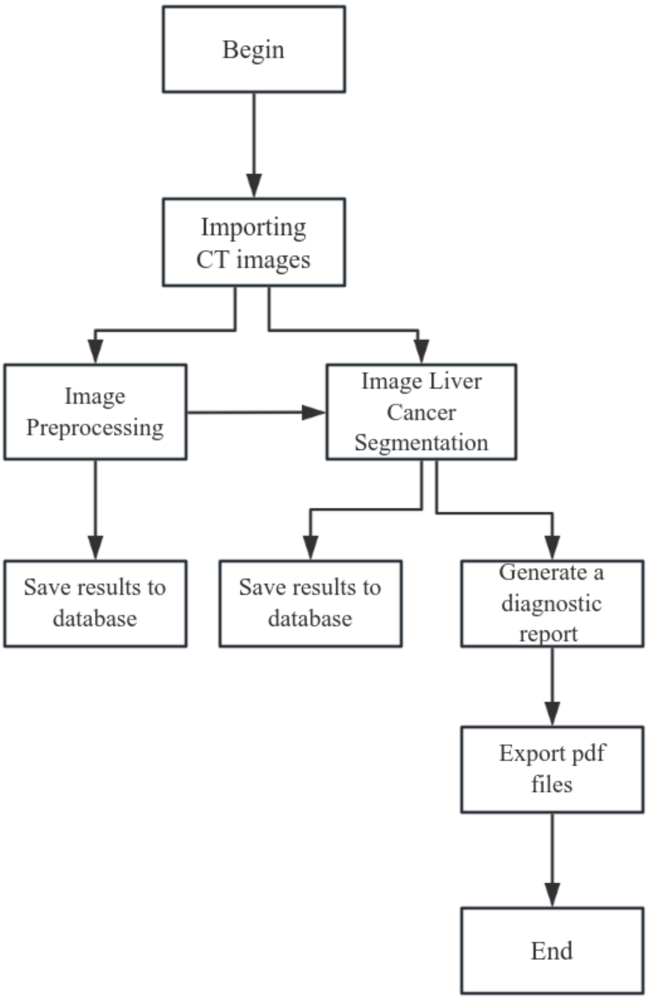


Fig.1: System Flowchart

2.1.1 Interface Design

This paper has designed a clear and straightforward interface for the liver preoperative diagnosis system, enabling doctors to get started quickly. There are functional button modules such as "Import Images", "Image Processing", and "Database Operation" in the menu bar at the top of the system. Through these functional modules, functions like CT image segmentation, saving results to the database, and generating preliminary diagnostic reports can be realized. The system is equipped with three canvases for display: the original CT image is shown on the left, the segmented results are displayed on the right, and the bottom is used to present preliminary diagnostic result information. This UI design facilitates doctors' intuitive observation and operation of the system functions.

2.1.2 Image Import Function

When using the system, doctors first need to import liver CT images and ensure they are in the correct format (bmp, jpg, png); otherwise, subsequent functional operations cannot be carried out. Clicking the "Import Images" button on the main interface triggers the bound clicked signal, and then a file selection interface pops up. At this time, import the correct CT images, and the imported images will be displayed on the interface for doctors to preview.

2.1.3 Image Preprocessing and Segmentation

After selecting CT images, doctors can perform corresponding processing operations on the images. As shown in Figure 2, they can click the "Image Processing" drop-down button to clearly see the distribution of various organs after CT image preprocessing, or directly click the "Image Segmentation" drop-down button to segment liver tumors directly.

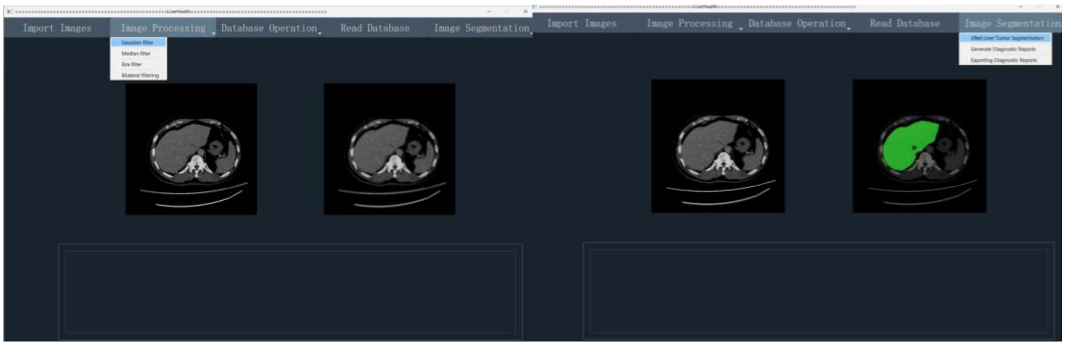
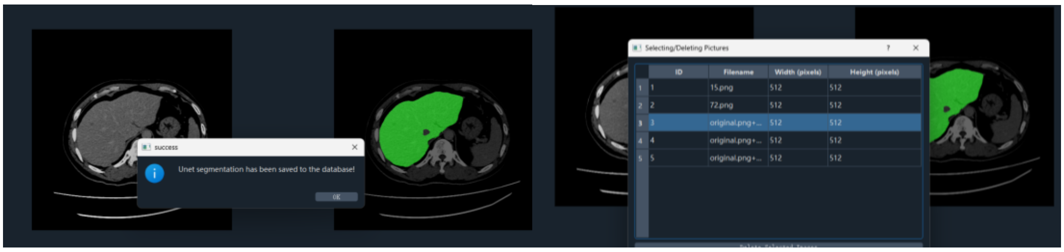


Fig.2: CT Image Processing and Segmentation

2.1.4 Database Saving and Reading

After the preprocessing step, doctors can choose whether to save the data to the database. As shown in Figure 3, clicking the drop-down button of "Database Operation" allows selecting whether to save according to the relevant processing results displayed on the right side of the interface. Then, clicking the "Read Database" button allows viewing the current saving status of the database, and corresponding data can be deleted as needed.



Unet segmentation has been saved to the database!

	ID	Filename	Width (pixels)	Height (pixels)
1	1	15.png	512	512
2	2	72.png	512	512
3	3	original.png+...	512	512
4	4	original.png+...	512	512
5	5	original.png+...	512	512

Fig.3: Saving and Reading Database

2.1.5 Generation of Diagnostic Reports

After liver tumor segmentation, doctors can click "Generate Diagnostic Reports" under the "Image Segmentation" drop-down button. As shown in Figure 4, the system can display the preliminary diagnosis content on the interface to assist doctors in auxiliary diagnosis.

```
The provided image is a segmentation mask of a liver, highlighting a region indicative of a hepatic
lesion. The following observations are made based on the analysis of the image:

1. **Lesion Identification:**
   - A distinct area within the liver has been segmented and highlighted in bright green.
   - The shape of the lesion is irregular, with an uneven border, which is often characteristic of
malignant tumors.
```

Fig.4: Diagnostic Results

2.2 System Testing

To verify the operability and interactivity of the system, this paper tests each process and function of the system, and the test results are shown in Table 1.

Table 1 System Testing

Serial Number	Test Function	Expected Result	Test Result
Test 1	Import Images	Selected images can be displayed on the interface	Passed
Test 2	Various Image Processing	Processing results are displayed on the right for comparison with the original image	Passed
Test 3	Liver Tumor Segmentation	Segmentation results can be marked with different colors and displayed on the right side of the interface	Passed
Test 4	Save to Database	A prompt pop-up appears after successful saving	Passed
Test 5	Read Database	A database pop-up appears to view saved results	Passed
Test 6	Generate Diagnostic Results	Diagnostic results are displayed at the bottom of the interface	Passed

3. Models

3.1 Liver Cancer Segmentation Models

The deep learning architectures employed in this system include Unet, Unet++, Unet3+, and Segformer. Weight files for these models are trained based on these architectures, then converted using the ONNX protocol to generate ONNX weight files. These files contain weight values, neural network flow information, input/output details of each layer, and other auxiliary data. The specific architectures and ONNX protocol are introduced as follows:

Unet, a convolutional neural network for image segmentation proposed by Olaf Ronneberger et al. in 2015 [10], primarily addresses medical image segmentation. Its structure comprises an encoder (gradually extracting features and reducing spatial resolution) and a decoder (restoring feature map size via upsampling to generate segmentation results). A key innovation is skip connections in the decoder, linking corresponding encoder and decoder feature maps to leverage multi-level features, enhancing segmentation accuracy and detail retention. It performs well on small datasets, with a simple, expandable structure.

Unet++ is an improved image segmentation network designed to boost performance through redesigned skip connections and multi-scale feature fusion [11]. Built on UNet, it enhances feature fusion and introduces deep supervision. It integrates embedded UNets of varying depths to improve segmentation of objects of different sizes (an advancement over UNet's fixed depth) and redesigns skip connections for flexible decoder fusion, breaking UNet's restrictive same-scale fusion to reduce the semantic gap between encoder and decoder, simplifying optimization.

UNet3+ draws on UNet and UNet++ [12], utilizing full-scale skip connections (directly combining high-level semantics and low-level details from multi-scale feature maps with necessary upsampling) and deep supervision (learning hierarchical representations from multi-scale aggregated feature maps). It captures full-scale fine-grained details and coarse-grained semantics. Each edge output connects to a hybrid loss function, aiding precise segmentation—especially for organs of varying scales in medical volumes. With fewer parameters than UNet++, it features deep supervision for efficient, parameter-light precise segmentation, plus a classification guidance module and hybrid loss function. It outperforms prior methods on liver and spleen datasets, highlighting organs with coherent boundaries.

Segformer, a semantic segmentation model combining Transformer and MLP, achieves efficient, accurate segmentation on multiple datasets via an innovative encoder and simplified decoder. This simple yet powerful method includes a hierarchical Transformer encoder (without position encoding) and a lightweight AllMLP decoder, avoiding complex designs for high efficiency and performance. It sets new state-of-the-art results on common datasets with strong zero-shot robustness. Its innovations include a position-encoding-free Transformer (adapting to arbitrary test resolutions) and a lightweight MLP decoder (aggregating multi-level information, combining local and global attention to enhance segmentation accuracy).

Open Neural Network Exchange (ONNX) is an open format for representing deep neural networks, supporting traditional machine learning models and unifying AI model exchange standards. It defines a platform/environment-independent standard, enabling AI model interoperability across frameworks and environments. Models trained in different frameworks can be converted to ONNX for easy deployment in compatible environments.

3.2 Introduction to Large Model

The large model called by this system is the Qwen-VL-Max multimodal large model. It uses a vision Transformer as the visual encoder to convert input images into feature vectors, compresses the image feature sequence using the visual language adapter of the cross-attention module, and adopts two-dimensional absolute position encoding to reduce the loss of position details. The language model part uses Qwen-7B for text processing and reasoning generation. In terms of training, the model is continuously optimized through three-stage training, including pre-training using large-scale weakly labeled data, multi-task pre-training, and supervised fine-tuning. In addition, the model also introduces a naive dynamic resolution mechanism, which can dynamically process images of different resolutions into different numbers of visual tokens. At the same time, it adopts multimodal rotary position embedding (M-RoPE), which decomposes position encoding into three components: time, height, and width, to better model the position information of multimodal inputs, and uses a unified image and video processing paradigm to enhance the versatility of the model. This system calls OpenAI's multimodal large model to understand images before and after liver tumorsegmentation, thereby generating diagnostic reports and realizing the function of auxiliary diagnosis.

4. Experiments

4.1 Experimental Dataset and Preprocessing

The system uses the public LIST dataset for liver tumors, in nii format, which is a CT scan image dataset focusing on liver and tumor segmentation. The dataset collects data from 7 different medical centers, including 131 training cases and 70 test cases. Each CT scan in the dataset is manually annotated, with annotations representing regions of interest in the form of binary masks [9]. In this paper, 28 cases are randomly selected from the dataset, and the dataset is divided into training set, validation set, and test set in the ratio of 8:1:1.

In the data preprocessing stage, since some of the dataset are full abdominal images but only the liver needs to be segmented, the redundant 'layers' without liver labels are removed. Firstly, the original CT images are subjected to window width and window level adjustment and cropping, limiting them to the range of [-200, 200], which helps to highlight soft tissues. Each pair of CT images and segmentation labels are read and converted into numpy arrays, retaining only the slices containing the liver. Secondly, the processed 3D medical images are converted into 2D image slices, and each image is transposed and rotated for data augmentation to prevent overfitting. Finally, the 2D image slices are divided into training set, validation set, and test set, and three text files are created to save the image and label paths of each dataset, preparing for later model training.

4.2 Experimental Environment and Evaluation Indicators

The model training in this paper is developed based on the PaddleSeg image segmentation suite of Baidu PaddlePaddle deep learning framework, and the experimental processing part is carried out on the PyCharm platform. Model training is performed efficiently using a Tesla V100 GPU with 32GB memory, and the number of iterations is set to 15,000 to ensure sufficient convergence of the model. During training, the initial learning rate is set to 0.01, and the batch size is set to 4.

In this model training, the following indicators are used to comprehensively evaluate and verify the performance of the 4 trained models in liver tumor segmentation tasks: Dice, mIou, Kappa, and Acc. The calculation formulas of each indicator are as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{1}$$

$$IoU = \frac{TP}{TP + FP + FN} \tag{2}$$

$$mIoU = \frac{1}{num\_classes}\sum_{i=1}^{num\_class} IoU_i \tag{3}$$

$$Dice = \frac{2\times TP}{2\times TP + FP + FN} \tag{4}$$

$$kappa = \frac{p_0 - p_e}{1 - p_e} \tag{5}$$

In the formula, FP (False Positive) refers to the result of incorrectly predicting other tissues or backgrounds as liver segmentation pixels; TP (True Positive) represents the prediction result correctly classified as liver segmentation pixels, that is, the pixels that the model predicts as liver and are actually liver; FN (False Negative) refers to the pixels that the model predicts as background but are actually liver, that is, incorrectly classifying liver pixels as background; TN (True Negative) refers to the pixels that the model predicts as background and are actually background, that is, the model correctly classifies background pixels; p_0 is the accuracy rate, and p_e is the probability of random consistency.

4.3 Experimental Environment and Evaluation Indicators

In the model comparison experiment, this paper uses the above 4 image segmentation indicators to test and evaluate the performance of the 4 models in Table 2 in liver tumor segmentation tasks. The experimental results are shown in Table 2:

Table 2 Model Performance Comparison

Models metric	mIou	Acc	Dice	Kappa
Unet	0.8909	0.9951	0.9321	0.9520
Unet++	0.8880	0.9935	0.9316	0.9504
Segfoemer	0.8896	0.9950	0.9336	0.9530
Unet3+	0.8377	0.9923	0.9040	0.9355

Experiments show that all indicators of the four models are close to 1, which generally indicates that these four models have excellent performance in liver tumor segmentation tasks.

MIou represents the average of the intersection-over-union for each class in the dataset. The closer it is to 1, the higher the degree of overlap between the predicted region and the real region. The mIou values of the above four models are all above 0.8, indicating that these four models will have relatively good segmentation results in applications.

Acc refers to the accuracy of the classifier, that is, the proportion of the total number of correctly predicted samples to the total number of predicted samples. The Acc values of the above four models are all above 0.9, which indicates that the four models have good performance and high prediction accuracy.

Dice is an indicator to measure the performance of segmentation algorithms, which is used to compare the similarity between the predicted segmentation results and the real labels. The Dice values of the above four models are all above 0.9, indicating that the predicted results are more similar to the real labels, that is, the segmentation results are more accurate.

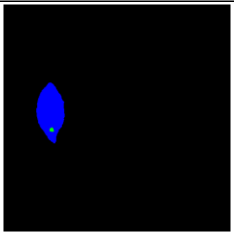
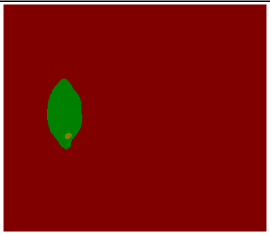
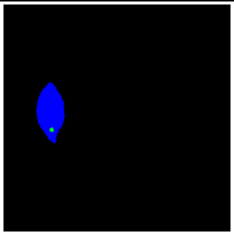
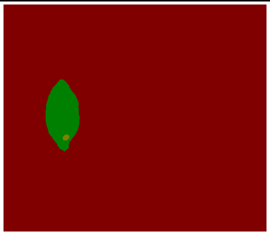
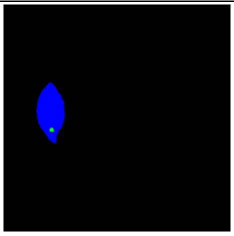
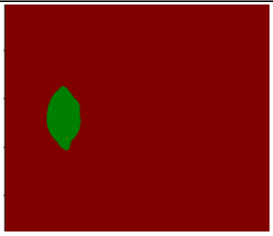
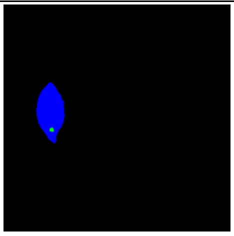
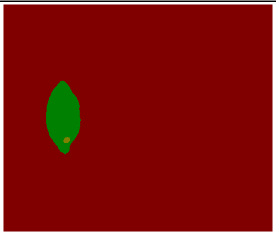
Kappa coefficient is an indicator for consistency test, which is usually used to measure the effect of classification. The Kappa coefficients of the above four models are all above 0.9, indicating a high consistency between the model's prediction results and the actual classification results.

Among them, Unet3+ is lower than the other three models in all indicators, and it also requires more GPU computing power and memory space in the actual training process. This also shows that compared with the other three models, Unet3+ shows lower performance in liver tumor segmentation tasks, and this model can be considered for use last in practical applications.

4.4 Model Comparison Test

To verify the effect of experimental data in practical application, this paper tests the liver tumor segmentation effect of the above 4 models, and the test results are shown in Table 3.

Table 3 Model Segmentation Result Test

Model	Original images of the test set	Test set label images	Test segmentation results
Unet			
Unet++			
Unet3+			
Segformer			

Through the above visual test comparison charts, it can be concluded that the 4 models selected in this paper can accurately segment liver tumors. However, due to the limitation of the experimental environment, the training of the Unet3+ model requires more than 32GB of memory and at least 5 hours of GPU computing power, resulting in the Unet3+ model failing to complete 15,000 iterations, failing to fully converge, and failing to accurately segment liver tumors. It can be seen from the test results of each model that except for the Unet3+ model, which requires large memory and computing power for training, the other 3 models can be fully trained with a memory capacity of 24GB or less.

5. Conclusion

Through multiple steps including data preprocessing, model construction, training and optimization, and evaluation of segmentation results, this paper constructs an efficient, robust, and highly generalizable auxiliary diagnostic system for liver tumor segmentation, providing reliable technical support for clinical auxiliary diagnosis and treatment. The core functions of the system include liver cancer tumor segmentation, generating auxiliary diagnostic reports, and saving result data to the database. The software system designed in this paper achieves good results in actual liver tumor segmentation, with all indicators above 0.8, and the segmentation models perform excellently. Meanwhile, the called multimodal large model conducts secondary evaluation on the preliminary segmentation results, helping doctors further carry out auxiliary diagnosis. Test results show that this system allows for convenient segmentation of the liver and tumors, enhancing the popularity and practicality of the technology.

Through the construction of the liver cancer auxiliary diagnostic system and the research on the performance of the 4 segmentation models, some issues that can be further studied are identified in this process.

(1) In this paper, liver tumor segmentation is realized by loading CT images into the segmentation model, and delays may occur in this process. Future work can explore the real-time performance and lightweight design of model segmentation to accelerate the segmentation speed of the model.

(2) This paper uses the Qwen-VL-Max multimodal large model to realize the auxiliary diagnosis of liver cancer, but the evaluation of this model is not specialized. Future work can consider researching this model to make it a large model dedicated to the medical field.

References

*[1] Wang Peiyu, Huang Qi, Wang Shaodong, et al. Key Points Interpretation of *Global Cancer Statistics 2022*[J]. Chinese Journal of Clinical Thoracic and Cardiovascular Surgery, 2024, 31(07): 933-954.*

[2] Zhao Haichao, Gao Qiang. New Progress in Liver Cancer Research and Diagnosis and Treatment in 2022[J]. Chinese Journal of Cancer, 2023, 33(04): 315-326.

[3] Cui Hui, Guo Yinghui, Cai Xiaohong, et al. Progress in the Application of Deep Learning in Liver and Liver Tumor Segmentation[J/OL]. Computer Engineering and Applications, 1-16[2025-06-02]. http: //kns.cnki.net/kcms/detail/11.2127.TP.20250328.1606.004.html.

[4] Hao Boqun. Design of Liver Image Segmentation System Based on Region Growing[J]. Journal of Tonghua Normal University, 2024, 45(10): 1-9.

[5] Xie Chang, Liu Changhong, Jiao Zhan. System Function Design and Software Implementation of Computer-Aided Method Research[J]. Information Recording Materials, 2023, 24(12): 201-204.

[6] Li Shufeng, Li Xiaolei, Wang Anshan. Design of Medical Image Automatic Diagnosis System Based on Deep Learning[J]. China New Technologies and Products, 2025, (01): 32-34.

[7] Xie E, Wang W, Yu Z, et al. SegFormer: Simple and efficient design for semantic segmentation with transformers[J]. Advances in Neural Information Processing Systems, 2021, 34: 12077-12090.

[8] Li Mingyang. Research on CT Image Segmentation Algorithm for Liver and Tumor Based on Semi-Supervised Learning[D]. Southwest University of Science and Technology, 2024.

[9] Wang Rongmiao. Research on Liver Segmentation Algorithm Based on Multi-Phase CT Images and Construction of Visualization Platform[D]. Soochow University, 2020.

[10] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. Springer International Publishing, 2015: 234-241.

[11] Zhou Z, Siddiquee M M R, Tajbakhsh N, et al. Unet++: Redesigning skip connections to exploit multiscale features in imagesegmentation[J]. IEEE Transactions on Medical Imaging, 2019, 39(6): 1856-1867.

[12] Huang H, Lin L, Tong R, et al. Unet 3+: A full-scale connected unet for medical image segmentation[C]//ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2020: 1055-1059.