Studies on Category Prediction of Ovarian Cancers Based on Magnetic Resonance Images

Master's thesis (Tech.)
University of Turku
Department of Future Technologies
2019
Yunfei Mao

Supervisors:
Assoc Prof. Tapio Pahikkala
Prof. Jinhua Yu

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

UNIVERSITY OF TURKU
Department of Future Technologies

MAO YUNFEI: Studies on Category Prediction of Ovarian Cancers Based on Magnetic Resonance Images
Master of Science in Technology Thesis, 50p
Computer Systems
Dec 2019

Ovarian cancer is the gynecological malignant tumor with low early diagnosis rate and high mortality. Ovarian epithelial cancer (OEC) is the most common subtype of ovarian cancer. Pathologically, OEC is divided into two subtypes: Type I and Type II. These two subtypes of OEC have different biological characteristics and treatment response. Therefore, it is important to accurately categorize these two groups of patients and provide the reference for clinicians in designing treatment plans.

In the current magnetic resonance (MR) examination, the diagnoses given by the radiologists are largely based on individual judgment and not sufficiently accurate. Because of the low accuracy of the results and the risk of suffering Type II OEC, most patients will undertake the fine-needle aspiration, which may cause harm to patients' bodies. Therefore, there is need for the method for OEC subtype classification based on MR images.

This thesis proposes the automatic diagnosis system of ovarian cancer based on the combination of deep learning and radiomics. The method utilizes four common useful sequences for ovarian cancer diagnosis: sagittal fat-suppressed T2WI (Sag-fs-T2WI), coronal T2WI (Cor-T2WI), axial T1WI (Axi-T1WI), and apparent diffusion coefficient map (ADC) to establish a multi-sequence diagnostic model. The system starts with the segmentation of the ovarian tumors, and then obtains the radiomic features from lesion parts together with the network features. Selected Features are used to build model to predict the malignancy of ovarian cancers, the subtype of OEC and the survival condition.

Bi-atten-ResUnet is proposed in this thesis as the segmentation model. The network is established on the basis of U-Net with adopting Residual block and non-local attention module. It preserves the classic encoder/decoder architecture in the U-Net network. The encoder part is reconstructed by the pretrained ResNet to make use of transfer learning knowledge, and bi-non-local attention modules are added to the decoder part on each level. The application of these techniques enhances the network's performance in segmentation tasks. The model achieves 0.918, 0.905, 0.831, and 0.820 Dice coefficient respectively in segmenting on four MR sequences.

After the segmentation work, the thesis proposes a diagnostic model with three steps: quantitative description feature extraction, feature selection, and establishment of prediction models. First, radiomic features and network features are obtained. Then

iterative sparse representation (ISR) method is adopted as the feature selection to reduce the redundancy and correlation. The selected features are used to establish a predictive model, and support vector machine (SVM) is used as the classifier.

The model achieves an AUC of 0.967 in distinguishing between benign and malignant ovarian tumors. For discriminating Type I and Type II OEC, the model yields an AUC of 0.823. In the survival prediction, patients categorized in high risk group are more likely to have poor prognosis with hazard ratio 4.169.

Keywords: Ovarian cancer; MRI; Radiomics; Segmentation;

Contents

1 Introduction

1.1 Overview of Ovarian Cancer

Ovarian cancer is the gynecological malignant tumor with low early diagnosis rate and high mortality, accounting for 22,500 deaths annually in China [1]. The cancer lacks specific clinical manifestations at the early stage and the course of disease progresses rapidly. 70%-80% of patients are diagnosed with advanced cancer with extensive metastasis in the basin and abdominal cavity at first doctor visit. The clinical treatment effect is not satisfactory. About 70% of ovarian cancer patients relapse within two years and the survival rate within two year is only 50% [2].

There are many subtypes of ovarian cancer with diverse classification method. Among them, ovarian epithelial cancer (OEC) is the most common gynecological malignancy, accounting for more than 70% of deaths [3]. Pathologically, OEC is divided into two subtypes: Type I and Type II. These two subtypes of OEC have different biological characteristics and treatment responses [3]. The overall survival of patients with Type I tumors was confirmed much higher compared with those with Type II tumors after 2 years of follow-up. Also, Type I OEC mostly presents in early stage, grows slowly and responds less frequently to platinum-based therapy [4]. On the other hand, Type II OEC, most are high-grade serous ovarian cancers (HGSC), is a highly invasive tumor with significant early sensitivity to platinum-based chemotherapy, leading to a poor prognosis [5]. To accurately classify patients into the right category may guide clinicians to design appropriate therapy plan.

At present, the diagnosis and treatment of ovarian cancer tumors can be generally divided into four steps [6]. The first step is the imagological diagnosis, radiologists make the preliminary diagnosis of the subtype of ovarian cancers based on the image characteristics of the tumor, such as size, texture and location. Magnetic resonance (MR) images of ovarian cancer are shown as Figure 1-1. In the second step, clinicians determine whether to perform fine-needle aspiration (FNA) on the patient with reference to the radiological diagnostic results and patients' clinical information. In

the third step, clinicians perform the FNA operation under the guidance of ultrasound, to obtain some tissues at safe position, and tests the obtained samples in pathological laboratory. In the fourth step, if the FNA results show the malignancy of the lesions, doctors perform the surgery to remove the ovarian tumors.

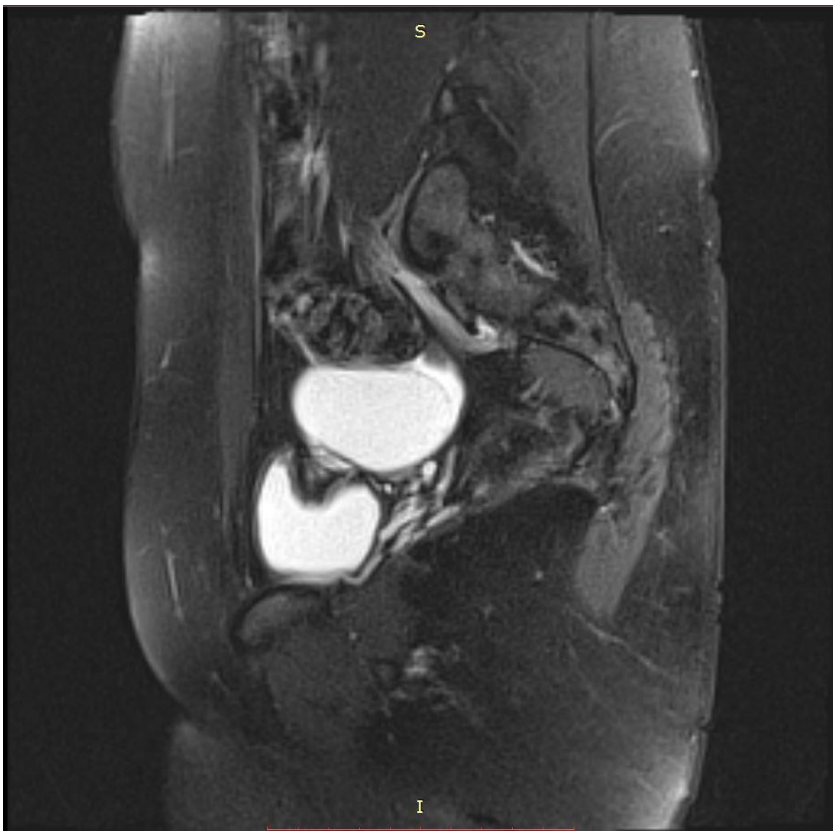

Figure 1-1 Ovarian Cancer Magnetic Resonance Image

Unfortunately, Type I and Type II do not have well recognized image characteristics. For this reason, during the imagological examination, the diagnosis given by the radiologist is largely based on individual judgment and not sufficiently accurate and informative.

Because of the low accuracy of the results of the preliminary imagological examination and the risk of suffering Type II OEC, most patients will undertake the fine-needle aspiration and tumor removal surgery in subsequent steps. However, both treatments have negative effects. Fine-needle aspiration may cause damage to the patient's body mechanism. Tumor removal surgery is even more risky and may result in paraneoplastic disease and nerve damage. Some patients with benign tumors or

Type I OEC, receive the over-treatment programs, which wastes medical resources and poses risks to patients' health [7]. More accurate classification techniques for ovarian tumors can reduce the possibility of under-treatment or over-treatment. At present, there is an urgent need for computer-aided method to assist radiologists in the diagnosis of ovarian cancer.

1.2 Medical Imaging Technology

Medical imaging technology provides the non-invasive acquisition of disease information for ovarian cancer. In the diagnosis of ovarian cancer, computed tomography (CT), ultrasound (US), and magnetic resonance imaging (MRI) are the three most commonly used medical imaging methods [8].

CT images reflect tissue characteristics via different absorption coefficients of X-rays in the human body [9]. CT has the advantages of short scanning time, fast localization of lesions and convenient imaging. However, the imaging procedure involves higher radiation, which may cause harm to the human body. Also, the resolution of CT images is lower than the magnetic resonance images.

Ultrasound is another imaging method in diagnosis of ovarian cancer. As ultrasonic waves propagate through the biological tissues, because of the heterogeneity of different kind of tissues, various wave propagation patterns occur, such as reflection, diffraction, and interference [10]. As the result, the property of received ultrasound signal illustrates the condition of the pelvic organs and can be used for further processing like imaging. The ovarian lesions are detected by identifying the differences and properties of the tissue in the ultrasound image. The ultrasound imaging is non-radiative, inexpensive and real time imaging approach. But the quality of picture generated is much worse than that of MR images. Therefore, the application of ultrasound imaging in the diagnosis of ovarian cancer is limited to some easy tasks, such as the classification of benign and malignant tumors.

The MR images are obtained by the interaction between the hydrogen nucleus and the electromagnetic pulse in the strong magnetic field. MRI provides excellent body

topographic structure and high soft tissue resolution with no radiation, which makes it suitable for observing pelvic conditions. [11]. By applying various emission strategy, such as the modification of the repetition time of the pulse, the echo time and the inversion time, different kinds of MRI sequences are acquired. Common magnetic resonance sequences include T1-Weighted Imaging (T1WI), T2-Weighted Imaging (T2WI), Fluid Attenuated Inversion Recovery (FAIR), T1 Contrast (T1C), Diffusion Weighted Imaging (DWI), and Apparent Diffusion Coefficient (ADC) [12]. Each magnetic resonance sequence has its own characteristics, specific for certain tissues.

In general, we can observe clearer anatomical structure of the pelvic and abdominal cavity of the patient from the MR image relative to the ultrasound image and the CT image. Also, the T2WI sequence of MRI images are designed to capture water area, which is especially useful for the radiologists to observe cancer lesions which is full of blood. Compared to the other two imaging techniques, MRI has more adjustable parameters and higher resolution. In summary, MR images are chosen as the research objects in this thesis.

Although MR images have so many advantages over other two imaging techniques, there still exists some difficulties for radiologists to accurately classify OEC. Therefore, it is important to discover the new effective algorithms based on MR images to help doctors diagnose the cancer.

1.3 Related Researches

The processing and analytic techniques in medical imaging have undergone tremendous development. Deep learning and machine learning algorithms are applied in this field to uncover deep characteristics in the image and capture features that are not easily noticeable to the naked eye [13]. With the help of these latest image models, better diagnosis and more precise prediction can be achieved.

Deep learning witnesses breakthroughs in natural language processing, computer vision, speech analysis, knowledge graph and so on since last decade, and has gained

widespread attention in recent years [14]. It is mainly supported by neural networks and is considered to be the most dynamic and promising research direction in all fields of artificial intelligence [15]. In 2012, with the integration of some new modules such as new activation functions and the compatible support for Graphics Processing Unit (GPU), deep learning achieved amazing results in ImageNet, a million-level image dataset [16].

Since then, deep learning has been successfully applied to various application due to its excellent performance, and medical image processing is one of the application scenarios [17]. Qiu et al proposed a multi-scale automatic segmentation of MR images based on CNN model, which classifies voxels into brain tissue classes [18]. Chlebus et al proposed a deep learning network where convolutional layers were adopted instead of a fully connected layers to speed up the segmentation process [19]. By using a cascading architecture, it connected the output of the first network with the input of subsequent networks to lessen computation. Bardou et al designed the new network structure with small kernels to classify each pixel in an MR image. The usage of the small kernel size reduces network parameters, which alleviated the risk of overfitting in building deeper networks. The authors also performs data enhancement and intensity normalization in the pre-processing step to facilitate the training process [20]. Rafiei proposed another CNN for brain tumor segmentation, in which he used the Dropout regularizer and Maxout activation function were adopted to handle overfitting problem [21]. In general, the deep learning program has been applied several times on magnetic resonance imaging, and the effectiveness of the algorithm has been well recognized.

However, there has been no deep learning study for MRI of ovarian cancer with deep learning. Recent studies for ovarian cancer are mainly based on machine learning or medical image processing, the method is often referred as radiomics.

Term radiomics is the combination of radiology and bioinformatics, and it is a new concept developed by these two disciplines [22] . Radiomics includes the quantitative extraction of high-throughput features from medical images, and the data mining from the image features. By establishing complex models, some targets such as

pathological information, molecular marker information, genetic information, and disease prognosis can be obtained. The concept of radiomics was initially proposed by Lambin in 2012 [23]. He first proposed the hypothesis and research procedure of radiomics, and applied image analytics technology to predict the relationship between gene expression and tumor characteristics and the relationship between image information and disease prognosis. Since then, radiomics began to become a research hotspot. Recently, radiomics has been used to analyze a variety of medical images and to provide different clinical information like tumor phenotypes and genetic protein status.

At present, the research on computer-aided diagnostic tasks for ovarian carcinoma is mainly based on radiomics approach. Kazerooni et al studied in the benign and malignant ovarian tumors classification over 55 patients by using descriptive features with linear discriminant analysis (LDA) classifiers [24]. Their study showed that the time-to-peak (TTP) and wash-in-rate (WIR) features showed a significant sensitivity in differentiating between malignancies and benign ovarian tumors. Rizzo et al researched on the relationship between radiomics features on CT images and prognostic factors in a sample of 101 patients with ovarian cancers. Their results showed that CT radiomic features were effective in predicting the risk of relapse in ovarian cancer cohorts during 24-month post-treatment [25]. Qiu et al compared two groups of CT ovarian images (pre-treatment and post-treatment). Their results showed that three key features: tumor volume, density, and density variance were useful in predicting 6-month progress-free survival (PFS). Their model achieved an accuracy of 0.831 in predicting PFS when applying the combination of these three features [26].

Although the previous studies made contribution in assisting radiologists in diagnosing ovarian tumors, there still exist some shortcomings in these researches. First of all, as far as we know, there has been no related work on OEC classification based on medical images. According to the previous introduction, the classification of Type 1 and Type II has significant clinical value and provides important reference for the design of treatment plans afterwards. Second, although deep learning has been applied to medical images of some diseases, currently there is no deep learning

application for ovarian cancer according to our knowledge. Deep learning has certain advantages over radiomics method. Deep learning has strong capacity to extract features from images, and it is especially useful for end-to-end model building. The previous research on ovarian cancer has not applied deep learning, which limits the performance of the models. Third, the previous model requires radiologists to manually segment the lesion area for the diagnosis of ovarian carcinomas. This operation consumes doctors' time, which hinders the promotion of the model in practical applications, and reduces the attractiveness of computer-aided diagnosis. Therefore, a fully automated diagnostic model for ovarian cancer is needed to optimize the workflow of the entire diagnosis and treatment, and to benefit both doctors and patients.

1.4 Objective and Highlight of the Research

The goal of this research is to combine the deep learning and radiomics to implement the automatic computer-aided diagnosis system of ovarian cancer based on MR cancer images to predict the category of OEC and fulfill other ovarian cancer diagnostic tasks. The model provides a diagnosis reference for the clinician, improving the efficiency and accuracy of the medical examinations.

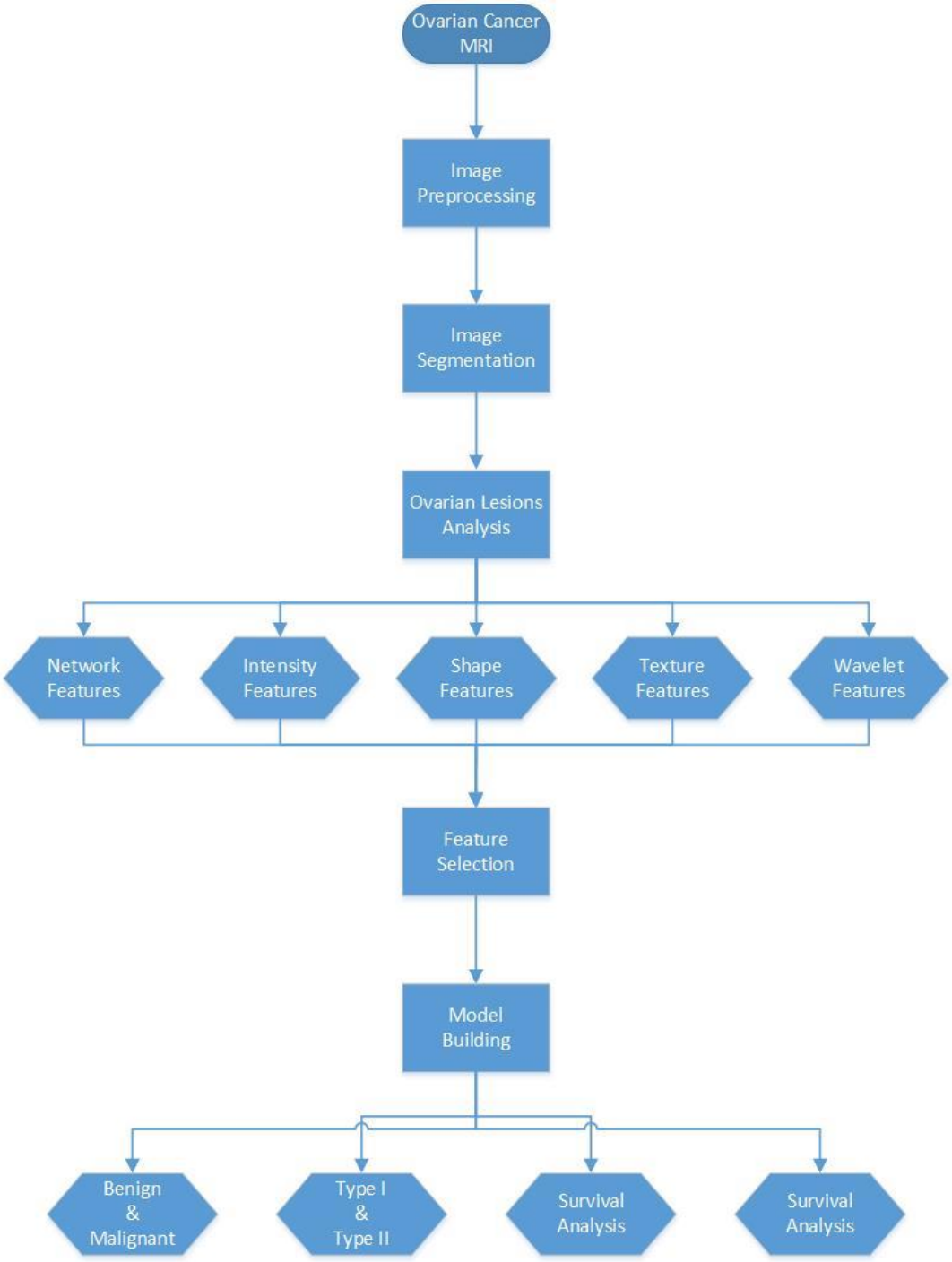

Figure 1-2: Flowchart of CAD of Ovarian Cancer

The overall framework of ovarian cancer diagnostic system in this thesis is shown in Figure 1-2. During the training process, only training data is included. Test set is used to obtain the performance of the final model. Also, the data augmentation operation in image preprocessing is omitted during the model evaluation.

The train model starts with the acquisition of the MR images of ovarian tumor, and

then the images are performed the preprocessing operation with data augmentation. The lesions are segmented through Bi-atten-ResUnet model, which is a revised U-net model proposed in this thesis. The details of the segmentation model will be introduced in the following chapters. After image segmentation work, we can identify the region of interest (ROI) of ovarian tumors. Then lesions analysis will be performed based on the segmentation results, which includes two parts. First, according to the radiomic features extraction system proposed in this thesis, radiomic features are obtained, including intensity, shape features, texture features, wavelet features. Second, the network features are acquired from the segmentation network to assist the lesion analysis. These two kinds of features are sent to perform the feature selection. Finally, the diagnostic tasks of ovarian cancers are fulfilled with the selected effective features. The diagnostic tasks include the classification between benign and malignant cancers, the classification between Type I and Type II OEC, the prediction of KI-67 and the survival analysis.

According to the research work, the main innovative points of this thesis can be summarized as follows:

1. This thesis proposes a segmentation algorithm for ovarian cancer lesions called Bi-atten-ResUnet based on MR images. The network is established on the basis of U-Net with adopting latest techniques such as Residual block and non-local attention module to deal with the problem of the complex background of ovarian cancer images. The work of the segmentation task is not only the prerequisite for the further analysis of ovarian cancer lesions. Meanwhile, the segmentation model has its own clinical value. By identifying the location of the lesion, it can help surgeons to complete the cancer removal operations more effectively and accurately. Therefore, the proposed segmentation model has certain significance for optimizing the diagnostic process of ovarian cancer.
2. In view of the problem that the current examination of ovarian cancer MRI is largely affected by subjective factors of radiologists; a quantitative description system of ovarian cancer images is proposed to quantify the information related to ovarian cancer. We acquire not only the radiomic features from the lesion area,

but also the network features of the segmentation network, which enhances the performance of our model.

3. The model proposed in this thesis applies multi-sequence MR images to analyze ovarian cancer. Previous ovarian cancer radiomics methods reported were mainly based on ultrasound images and CT images. The MR images used in this thesis have higher resolution and are more suitable for the diagnosis of ovarian cancer. Also, the results of multi-sequence and single-sequence prediction models are compared, and the validity of multi-sequence method is proved.
4. According to our studied samples, the proposed diagnosis system of ovarian cancers achieves good performance in benign and malignant ovarian tumor classification, Type I OEC and Type II OEC classification tasks. The accuracy of the model prediction exceeds the diagnostic results of the radiologist, which illustrates that our model can provide the effective reference for radiologists when they interpret the MR images diagnosis.

1.5 Datasets

The data collection work in this research was approved by the review board of Gynecological and Obstetric Hospital, School of Medicine, Fudan University. First, altogether 438 patients with suspected with ovarian tumors were retrieved from the hospital Picture Archiving and Communication System (PACS, GE). Then, we set the inclusion criteria for select the appropriate candidates for the study: 1) no previous pelvic surgery; 2) no previous gynecological disease history. Finally, the total studied samples we acquired were 280 patients (Table 1-1 and Figure 1-3). The class labels are based on the pathology reports. Pathological examination is the golden standard in the ovarian cancer diagnosis.

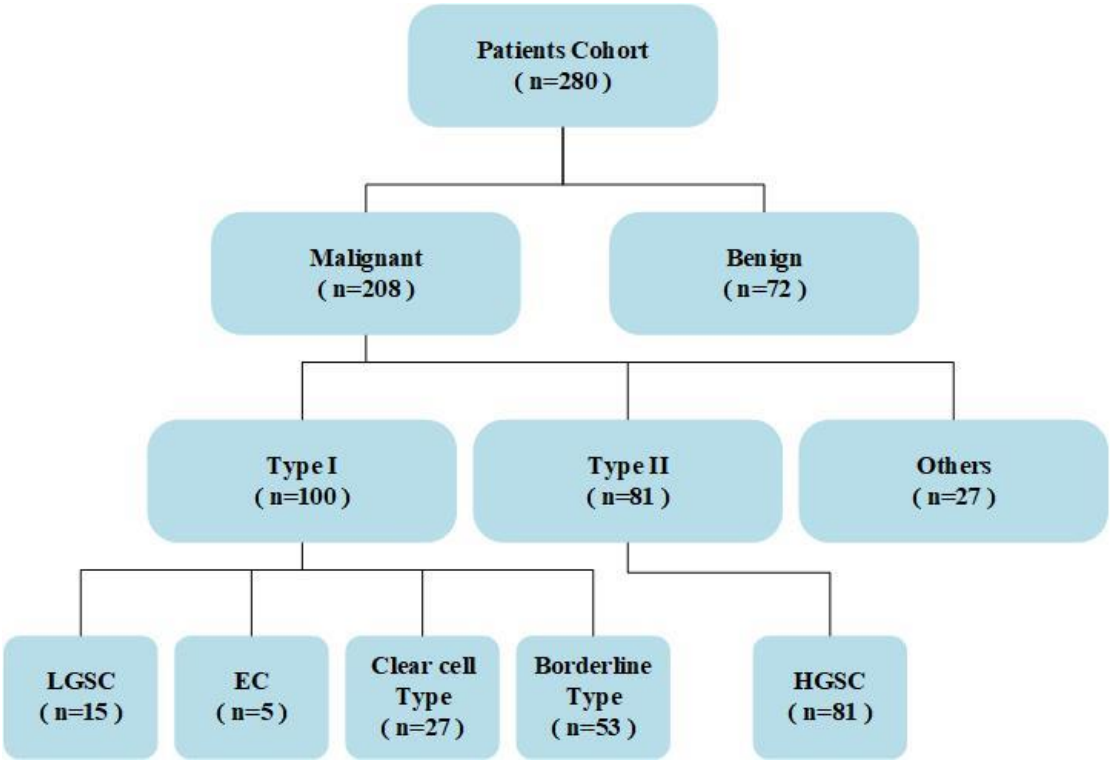

Figure 1-3: Flow chart of patient selection for different classification tasks

Table 1-1. The summary of the pathological types of the selected samples.

Pathological type	Numbers	Age(yrs.)*
Type I ovarian cancer	100	43.3 ± 13.9
Type II ovarian cancer	81	54.5 ± 9.9
Others malignancies	27	43.6 ± 15.3
Benign etiologies	72	43.6 ± 19.5
Total	280	46.2 ± 11.8

The whole data is divided into two parts: a training set and an independent testing set. In terms of age, CA-125, Ki67 expression and category information, there is no statistically significant difference between training cohort and independent testing cohort (Table 1-2).

Table 1-2. Clinical and pathological data summaries in training and testing cohort.

		Training cohorts (N = 195)	Testing cohorts (N = 85)	P value
Age(yrs.)		47.08±15.81	44.66±16.26	0.884
	<30	35 (17.9%)	30 (35.3%)	
	30-50	66 (33.8%)	18 (21.1%)	
	>50	94 (48.3%)	37 (43.6%)	
Ki-67 expression (%)		28.28±23.20	22.72±21.91	0.073
	<50	139 (71.3%)	70 (82.4%)	
	50-75	46 (23.6%)	8 (9.4%)	
	>75	10 (5.1%)	7 (8.2%)	
CA-125 level (IU/L)		483.69±821.00	513.04±1105.27	0.555
	<35	28 (14.4%)	25 (29.4%)	
	35-200	72 (36.9%)	32 (37.6%)	
	200-500	35 (17.9%)	15 (17.6%)	
	>500	60 (30.8%)	13 (15.4%)	
Category				0.113
	Type 1	60 (30.8%)	37 (47.1%)	
	LGSC	9 (4.7%)	6 (7.1%)	
	EC	1 (0.5%)	4 (4.7%)	
	Clear Cell type	17 (8.7%)	9 (11.8%)	
	Borderline type	33 (16.9%)	18 (23.5%)	
	Type 2			
	HGSC	63 (32.3%)	17 (21.2%)	
	Benign	52 (26.7%)	23 (23.5%)	
	Malignant	20 (10.3%)	8 (8.2%)	
FIGO*				0.111
	IA	44 (35.7%)	26 (44.8%)	
	IIA	14 (11.4%)	2 (3.4%)	
	III	55 (44.7%)	27 (46.5%)	
	IV	10 (8.1%)	3 (5.2%)	

MR images included in the thesis were acquired via a 1.5-T MR system. Multi-MRI sequences used for the assessment in this thesis include the axial T1-weighted imaging (Axi-T1WI), sagittal fat-suppressed T2-weighted imaging (Sag-fs-T2WI), and coronal TSE T2WI (Cor-T2WI). These four sequences of images are considered most useful when radiologists diagnose ovarian tumors. Based on the prior knowledge,

we assume that they are also the suitable samples for image processing. Manual annotations of ovarian lesions were performed by one experienced radiologist (H.Z. with more than 10 years of experiences) on the MATLAB platform.

1.6 Thesis Structure and Organization

The content of this article is organized into six chapters.

Chapter one mainly introduces the importance of ovarian cancer diagnosis based on MRI, the significance of ovarian cancer lesion segmentation and the ovarian cancer subtype prediction. The chapter also briefly describes the advantages of radiomics and deep learning in medical image segmentation and classification tasks. Finally, the research objective and structure of the thesis are included.

Chapter two introduces the background knowledge and related research work of deep learning, including convolution layer, pooling layer, loss function and so on. These cutting-edge technologies provide the necessary support for the research in this thesis.

Chapter three proposes a segmentation algorithm for the segmentation of ovarian cancer lesions in MR images. According to the characteristics of the lesion image, the base network structure of U-Net is selected. Some latest deep learning techniques, such as non-local attention and Residual blocks are adopted to improve the segmentation performance of the model. The effectiveness of the improved algorithm is evaluated by some comparison experiments. The output of the image segmentation part will be used to guide the diagnostic model in the next chapter.

Chapter four explores the model for predicting ovarian tumors subtypes and other tasks related to ovarian tumors. The combination between radiomic features, including intensity features, texture features, shape features, and deep learning network features are fulfilled. Combined with the segmentation algorithm introduced in Chapter three, a fully automatic ovarian cancer diagnostic model without annotation prerequisite is proposed to provide a reference for radiologists when they make diagnosis.

Chapter five is the conclusion part. The contribution of this thesis is summarized, the

shortcomings of the model are analyzed and the plan for model further optimization is forecasted.

2 Introduction of Convolutional Neural Network

2.1 Introduction

Chapter two mainly introduces some core knowledge of deep learning, which is the basis of the model proposed in this thesis. The convolutional neural network (CNN) structure is the foundation of the network architecture used in this thesis. Some content of deep learning will be introduced in the next paragraphs.

2.2 Convolutional Neural Network

CNN originated from the feedforward artificial neural network (ANN). Because of the characteristics of shared weight architecture and translation invariance, CNN is regarded as shift-invariant or spatially invariant artificial neural network. Therefore, it is the most popular network in solving the computer vision tasks. CNN is mainly inspired by biological processes, and the connection between neurons is similar to that in animal visual cortical tissue. Each visual neuron is responsible for responding to a small area in the visual field. The visual system integrates the signals received by massive neurons to obtain the overall content of the field of view. Compared with other image processing algorithms, CNN is considered to have low dependence on preprocessing. In many cases, even the original image can be directly used as a model input, which is very convenient to use. When CNN model is used, there is no need to manually design special features, and the model can independently learn the characteristics from a large amount of sample data. CNN applies convolution and pooling operations, which is highly computationally efficient and allows the CNN model to run on some devices with low computing power. Currently, CNN is widely used in medical images and natural images, and is considered to be one of the most mature ANNs. In this thesis, CNN is mainly used for lesion segmentation. In the

following paragraph, its core modules and algorithms will be introduced.

2.2.1 Convolutional Layer

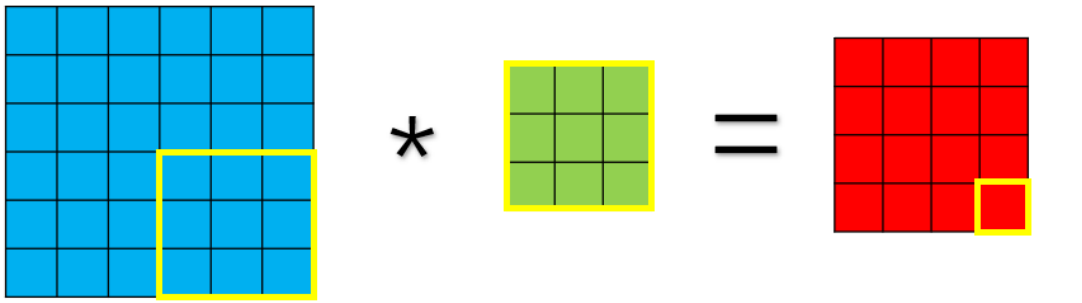

Figure 2-1: Convolutional calculation method

The convolutional layer is the core part of CNN. It consists of a set of learnable filters or kernels that can perform most feature extraction work. Each convolutional kernel is four-dimensional spatially, namely length, width, input depth, and output depth. The sizes of length and width are usually small odd numbers, such as 3*3 or 5*5, to limit the computation burden within convolutional layers. During forward propagation period, each filter slides over the input vectors, making the convolutional calculation. The calculation actually is the dot product between the kernel and the corresponding positions of the input feature map. When the filter slides over the all the candidate positions of the input vector, one two-dimensional output feature map is generated which is the response of the filter at each spatial position. There are a set of filters in each convolutional layer, and each filter will produce one individual two-dimensional activation map. Finally, all the maps are stacked along the depth dimension and the output is generated. The convolution operation follows equation (2-1).

$$B(i,j)=\sum_{m=0}\sum_{n=0}K(m,n)*A(i-m,j-n) \tag{2-1}$$

Where K symbolizes the convolution kernel, and A is the input feature map and B is the output feature map.

The main reason for introducing convolutional layer is to reduce the number of model

parameters and pay more attention to the local relationship. The characteristics of the convolutional layer just meet these two requirements.

(1) Weight Sharing:

CNN is often used to deal with the image data, the very high dimensional data. In this condition, it is impractical to connect neurons to all neurons in the previous layers, which leads to unacceptable heavy computation and huge model parameters. During convolutional layers forward, each time the same convolution kernel is used to translate the entire feature image to some specific information, which largely reduces the number of parameters and the computational operations.

(2) Local perception:

It is generally considered that the pixels of the image have spatial correlation, the closer the distance is, the closer relationship is. Convolutional kernel, with much smaller size compared to the whole feature map, focusing on every local area across the map, extracting the characteristics in the certain region.

Each convolutional kernel represents the knowledge which the model learns from the sample data. For example, kernels in the first layer may learn to analyze the edges in certain directions or spots of a certain color and kernels in higher layers learn the entire object outlines or shape patterns.

2.2.2 Pooling Layer

After the features are extracted through the convolutional layer, it is desirable to use these features for further processing. However, if all the obtained features are used for analysis, it will lead heavy calculation and over-fitting problems. Therefore, it is necessary to use the pooling operation to acquire the most representative features, thereby simplifying the calculation and improving the robustness of the model. The features after the pooling layer discard the insignificant details and retain only useful knowledge that are important to the results. Also, the exact position of the feature in the whole image is blurred, and the relative position between the features is preserved, so that the influence of the distortion of the object shapes can be well overcome.

The pooling process applies a down-sampling function without any parameters. A window slides with a specific step size from the upper left corner to the lower right corner of the feature map and calculates and output map of the corresponding area. The most common size of the window is 2, which reduces the map by 75%, as is illustrated in Figure 2-2. The most common way to pool is maxpooling.

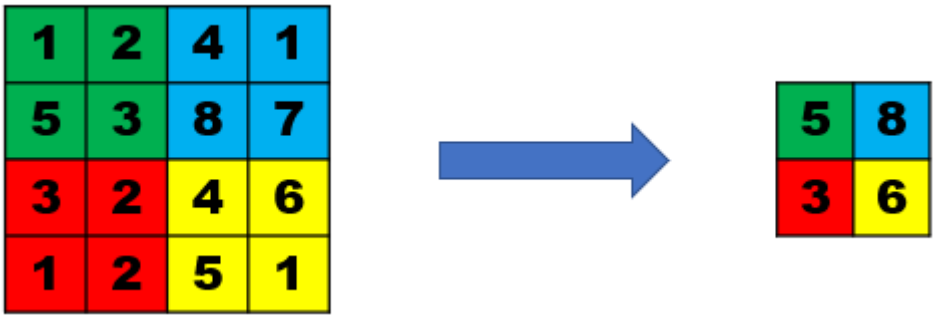

Figure 2-2: Maxpooling calculation method

2.2.3 Transposed Convolutional Layer

Transposed convolution, also known as deconvolution, generally used for the upsampling operation to restore the low-resolution images to the size before pooling. As shown in Figure 2-3, the operation of deconvolution is similar to convolution, and actually the calculation can be based on the convolution operation. A kernel with certain size is used to filter the input image, and the output result will be larger than the input image. The stride size determines the ratio between output size and input size.

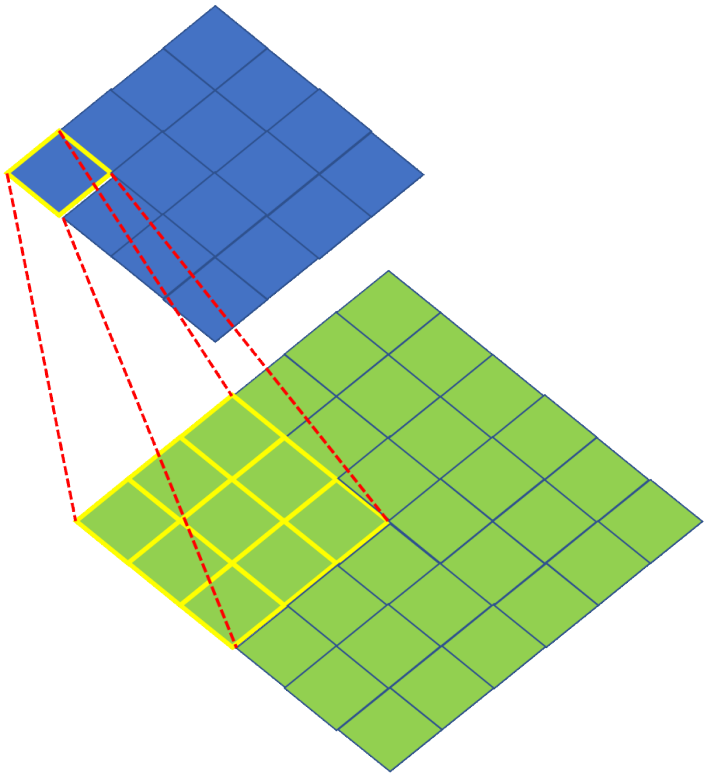

Figure 2-2: Transposed convolution calculation method

The stride in each transposed convolution layer is set two in this thesis, which amplifies the width and height of the input feature map two times. The reason is that the pooling operation performed in the shrinking path reduce the width and height of feature map by 2, so that the final output of the model will be the exact same size of the original input images.

2.2.4 Loss Function

The loss function is the key to the effectiveness of the deep learning model, which plays an important role in the training process of the model. The loss function measures the distance between the prediction of the model and the true labels, so as to evaluate the fitting ability of the network. The larger the loss value means the worse the network's ability to fit, or the larger the bias of the model. The loss function is the entry for the back propagation.

In CNN classification model, the cross entropy (CE) loss function is one of the most

commonly used loss functions of the researcher. It is defined as follows:

$$\mathrm{E}(t, y) = -\sum_{j} t_j \log y_j \tag{2-2}$$

Where t and y respectively represent the real category label of the training sample and the predicted value of the CNN model, and y_j represents the each output result of the Softmax function, as is shown in formula (2-3)

$$\mathrm{y}_j = \mathrm{softmax}(z_j) = \frac{e^{z_j}}{\sum_{j} e^{z_j}} \tag{2-3}$$

In image semantic segmentation task, the pixel-level cross entropy is the most common used loss function. In this case, the unit of classification is each pixel rather than an object. There are two categories of the labels: image foreground and background. Usually, the loss function compares the target vector encoded by the one-hot code with the prediction vector to obtain the average accuracy of all predictions. Since the pixel-level cross entropy calculates the influence of each pixel equally, the loss value will be dominated by the class with largest numbers of pixels in the training process when the sample category is imbalanced. And it is not suitable for the segmentation tasks when the border of object changes softly.

Another popular loss function for image segmentation tasks is based on the Dice coefficient, which essentially measures the overlap degree of two object. The value of Dice coefficient ranges from 0 to 1 and Dice coefficient of 1 represents completely overlapping of the predictions and true labels. Formula of Dice coefficient is as follows:

$$\mathrm{Dice} = \frac{|A \cap B|}{|A| + |B|} \tag{2-4}$$

Where $|A \cap B|$ represents the total number of common elements in the A and B sets. To calculate numerator $|A \cap B|$, we multiply the prediction mask and the label mask, and sum up the matrix result of the multiplication to achieve $|A \cap B|$. The numerator of the dice loss function focuses on the common activations between the

prediction and the target mask, and the denominator is related to the amount of activations in each mask. This characteristic can be used to normalizes the loss based on the size of the target mask, so it is easier for the model to learn from classes that are less distributed in the image.

Therefore, the adoption of the Dice loss function can be used as the complement to the pixel-level cross entropy to alleviate the problem caused by imbalanced distribution of sample categories.

3 Segmentation of Ovarian Tumors

3.1 Introduction

Image segmentation of ovarian cancer lesions is a crucial step in the automatic diagnosis system for ovarian cancer. The purpose is to separate the tumor lesion area from the normal organs in the MR image. After that, the lesions can be analyzed based on the segmented tumor area to figure out some specific characteristics and to achieve diagnostic tasks such as tumor subtype prediction and survival analysis, which provides a reference for assisting doctors to make more accurate diagnosis [17], [27]. If the lesion is not accurately segmented, further analysis of the lesion image is not accurate, and the final prediction result will greatly deviate from the real situation. The segmentation work itself also has important clinical value. Accurate positioning of the tumor area can be used to provide a reference for clinicians for operative preparation. At present, there have been studies on the segmentation of different lesions on different MR images, but the automatic segmentation of ovarian tumor has not been covered before. This section will present an ovarian lesions segmentation algorithm based on combination of image processing, deep learning and post-processing method.

3.2 Image Preprocessing

Image preprocessing in this thesis is mainly the data augmentation operation. The target of the data augmentation is to solve the over-fitting problem. Over-fitting is a common problem in computer vision tasks, and the essential reason is that there are too many model parameters with too little data. The simple and intuitive approach to mitigate overfitting is to augment the number of training data. However, in practical situations, the cost of obtaining true samples is high, and it is often difficult to directly augment the training sets. The condition is especially obvious in medical imaging

problems because the source of data is heavily limited by the number of patients attending the hospital, and each MR examination costs a lot. The usage of patient medical data is also subject to a lot of regulation. Fortunately, data augmentation of image samples is relatively easy compare to other kinds of data format. For the original image in the training set, we can generate a new sample by simply translating or rotating it. Although the new image may generate some irrelevant sample features, the corresponding structural features of the original image are consistent. In this thesis, each image sample will randomly generate 4 images based on the random combinations of augmentation approaches. Image augmentation method include flipping, rotating, translating, brightness changes, contrast changes, etc.

3.3 Segmentation Network

3.3.1 Network Architecture

The architecture of segmentation network Bi-atten-ResUnet proposed in this thesis is as shown in Figure 3-1. The name of the model shows that it is established on the basis of U-Net with Residual block and non-local attention module. It preserves the classic encoder/decoder architecture in the U-Net network, as well as the skip connections. The encoder part is reconstructed by the pretrained ResNet to make use of transfer learning knowledge, and bi-non-local attention modules are added to the decoder part on each level. The application of these techniques enhances the network's performance in segmentation tasks. Next, three core techniques: U-Net, Residual block and non-local attention mechanism will be introduced.

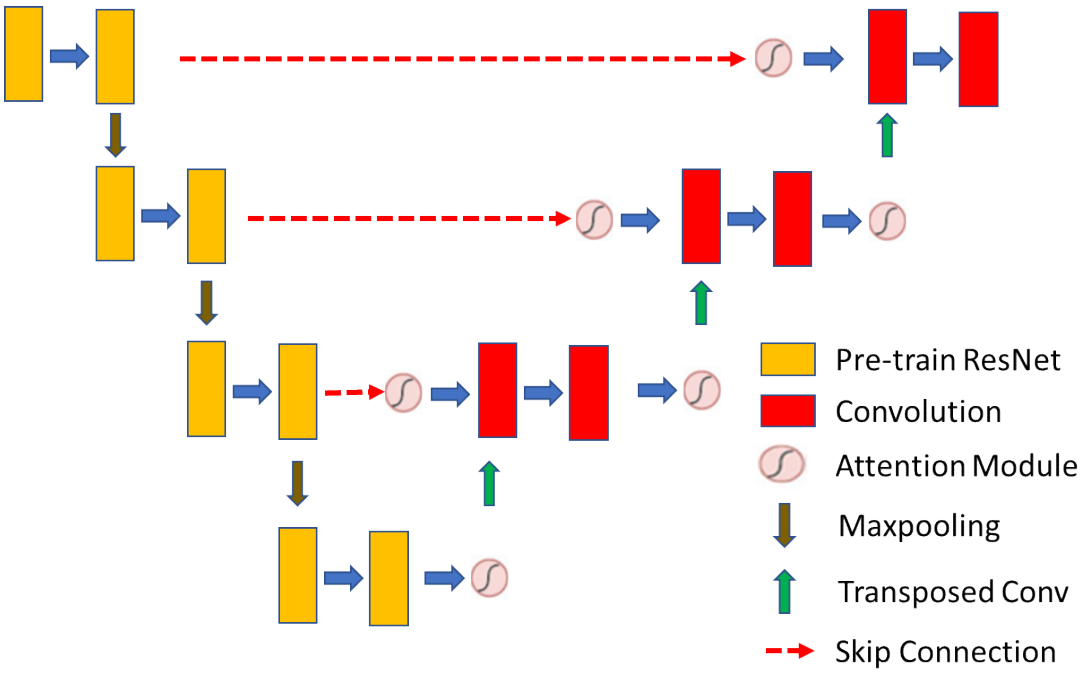

Figure 3-1 Segmentation Neural Network Architecture

3.3.2 U-Net

U-Net [28] was proposed in 2015 and has been widely used in many fields of industry, such as medical image segmentation and component image segmentation. Like many segmentation networks, U-Net uses an encoder/decoder structure. The skip connection is used to connect the low-level features with the high-level features to form the combination of texture information and semantic information. This architecture helps the decoder to better restore the detailed features of the picture. It also makes the encoder and decoder no longer two separate processes to improve prediction accuracy. The decoder part of the classic U-Net is different from the FCN [29]. When dealing with complex background image data, U-Net shows better segmentation result than the FCN. Ovarian cancer MR images usually contains lots of texture details, so U-Net is chosen as the backbone network in this thesis.

The encoder/decoder structure is actually not a fixed form, but an architectural idea. We can adjust the number of encoders or decoders blocks as needed, or even completely change its original forms. In this thesis, the entire encoder is reconstructed

by the ResNet which is pre-trained on imagenet datasets, and the attention module is added in the decoder part.

3.3.3 Residual Module

ResNet improves the performance of the network significantly by simply adding a part of the shortcut connection, compared with the conventional stacked CNN [30]. The existence of shortcut makes the ResNet block provide a highway to pass "identity mapping" information. With such unique design, ResNet module is considered to be one of the most effective solutions to combat the vanishing gradient problem. [31], [32]. The structure of the ResNet block is as shown in Figure 3-2.

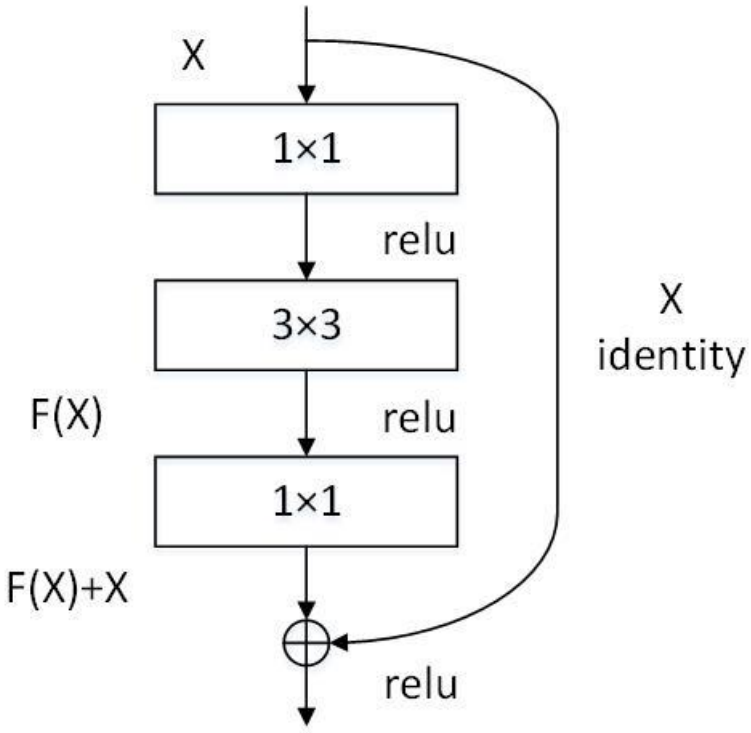

Figure 3-2 Residual Module

The reason for the vanishing gradient is that in deep CNN network, the gradient propagation from the bottom layer to the top layer needs continuous multiplication, and it may result in the small gradient, which makes the deep network model not easy to train [33]. An extreme solution to avoid the vanishing gradient is to let the added

layer not learn new features at all, only to replicate the characteristics of the shallow network. These layers are identity mapping, which is the essential idea of ResNet. In this condition, the residual is 0 and the performance of the deep network should be at least as good as that of the shallow network. This situation prevents the performance of the network from degrading. In actual cases, the residual is not 0, which makes the network layer can at least learn some new information based on the input. ResNet mitigates the training degradation of deep networks through residual learning, so that deeper networks can be trained. The formula for the residual unit can be expressed as shown in equation (3-1):

$$H(X) = F(X) + X \tag{3-1}$$

3.3.4 Non-local Attention Module

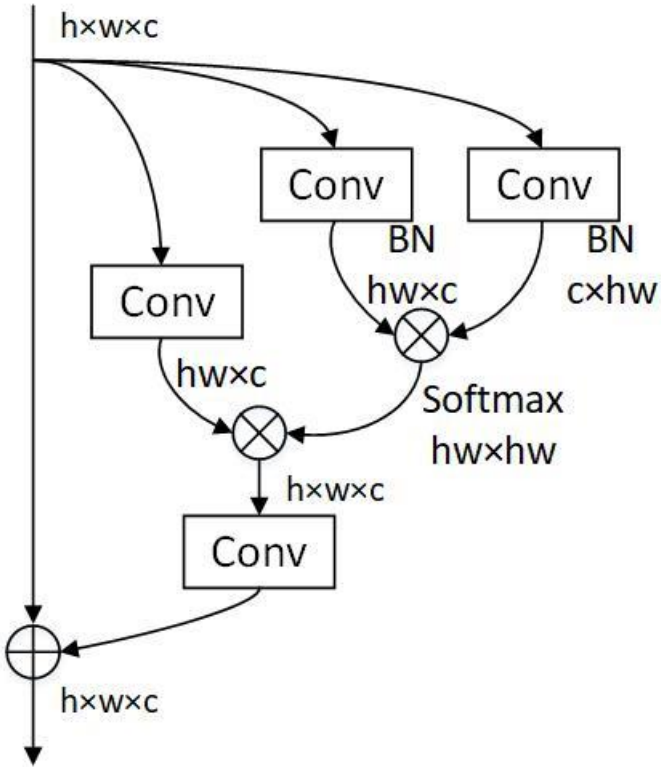

Figure 3-3 Non-local Attention Module

The idea of non-local attention module originates from the way we observe things.

When people observe objects, we don't pay attention to the whole part of the object. Instead, we are generally concerned with only a certain part. In other words, the attention distribution of each position is different. The concept of attention mechanism was first proposed in 2014 [34]. Up to now, attention-based models have been widely used in a variety of deep learning tasks. Wang et al. proposed a non-local attention module [35], as shown in Figure 3-3. The source of the idea is that the convolution unit in CNN only pays attention to the area of the neighborhood kernel size at a time. Even if the late receptive field is getting larger, it is still a local area operation, thus ignoring the other regions of the feature maps, such as distant pixels' contribution to the current region. Non-local attention module solves the problem by utilizing all the pixels in the feature map to calculate the attention coefficient of each pixel.

The purpose of applying the attention module in this thesis is to let the network focus on the tumor area. In addition, we also compared the performance difference between the single attention module and bi-attention module.

3.4 Post-Processing

On the basis of the neural network model, further post-processing operations are made to enhance our segmentation results. The post-processing operations are mainly divided into two steps. The first step is called connected domain area filter (CDAF). As a result of network segmentation, some small misdetected targets are sometimes generated.

Figure 3-4 Connected Domain Area Filter

As shown in Figure 3-4, these small connected domains usually have much smaller number of pixels compare to the true tumor areas, so they can be removed by CDAF. The minimum area threshold is set in CDAF, the connected domains whose areas are less than the threshold will be eliminated from the target predictions. Actually, in the experiment, this threshold is set to 40.

The second step is called location judgment (LJ). When observing the segmentation results, we find that the neural networks sometimes misdetect bladders as ovarian tumors. The reason can be that the bladder is rich in water, so it shows a high signal in the MR images. This pattern is similar to the ovarian cancer tumor, since the tumor contains a lot of blood, also resulting in high activations in the images. Therefore, the model is easy to misjudge the bladder as a tumor, which will undoubtedly have the bad influence on the diagnosis accuracy. At the same time, the CDAF does not work in this case due to the large area of the bladder. Therefore, we propose another post-processing method based on physiological knowledge. Since the position of the bladder is always lower than the uterus and ovaries in the pelvic cavity, LJ can be used to remove the bladder with lower location and retain the correct ovarian tumor. The process of this operation is shown in Figure 3-5.

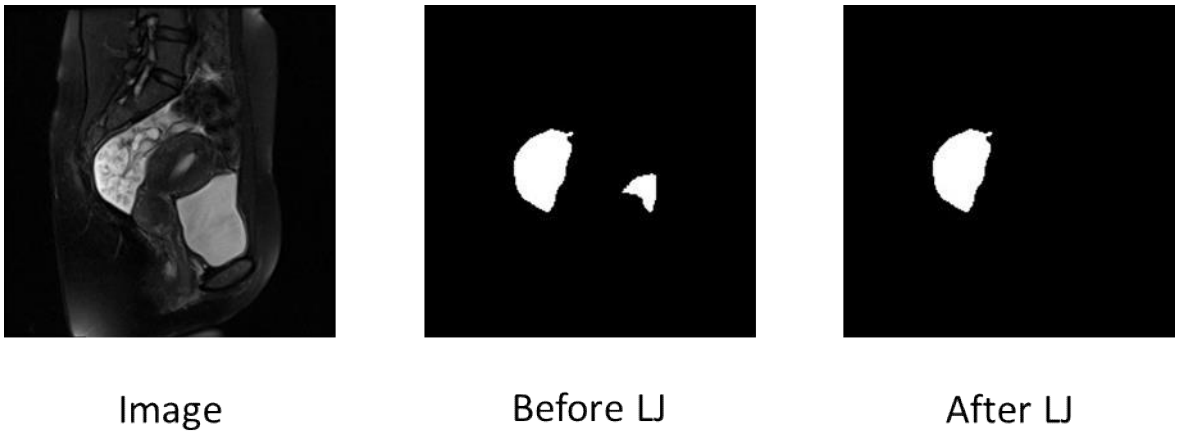

Figure 3-5 Location Judgment

3.5 Result

In this section, the performance of final proposed method Bi-atten-ResUnet and the partial improved form is compared. The structural differences between the different models and the corresponding performances are shown in Table 3-1. There are also experiments to compare the results with and without postprocessing. The split of training and testing data is as shown in Table 1-2. The metrics used for evaluation in this thesis is the Dice coefficient. The theory and calculation of the Dice coefficient has been introduced in the loss function section in Chapter 2. The larger the value of the Dice coefficient, the closer the prediction result of the model is to the real value, and the better performance of the model.

Table 3-1. Results of Different Segmentation Models

Method	Modal	Pretrained ResNet	Non-Local Attention	Bi-Non-Local Attention	Post-Process	Dice
U-Net	Sag-fs-T2WI				√	0.817
ResUnet	Sag-fs-T2WI	√			√	0.889
Atten-ResUnet	Sag-fs-T2WI	√	√		√	0.906
Bi-Atten-ResU	Sag-fs-	√		√		0.885

net no Post-procesing	T2WI					
Bi-Atten-ResUnet	Sag-fs-T2WI	√		√	√	**0.918**
U-Net	Cor-T2WI				√	0.834
ResUnet	Cor-T2WI	√			√	0.890
Atten-ResUnet	Cor-T2WI	√	√		√	0.901
Bi-Atten-ResUnet no Post-procesing	Cor-T2WI	√		√		0.863
Bi-Atten-ResUnet	Cor-T2WI	√		√	√	**0.905**
U-Net	Axi-T1WI				√	0.785
ResUnet	Axi-T1WI	√			√	0.825
Atten-ResUnet	Axi-T1WI	√	√		√	**0.837**
Bi-Atten-ResUnet no Post-procesing	Axi-T1WI	√		√		0.810
Bi-Atten-ResUnet	Axi-T1WI	√		√	√	0.829
U-Net	ADC				√	0.772
ResUnet	ADC	√			√	0.804
Atten-ResUnet	ADC	√	√		√	0.815
Bi-Atten-ResUnet no Post-procesing	ADC	√		√		**0.820**
Bi-Atten-ResUnet	ADC	√		√	√	**0.820**

According to Table 3-1. The final model Bi-atten-ResUnet with post-processing achieves the best performance. Each deep learning techniques have contributed to some improvements of the model. Among them, the pre-trained Residual block gets maximum increase, which may prove why ResNet is so popular in all kinds of image tasks currently. Post-processing also brings certain enhancement.

3.6 Summary

In this chapter, we apply the combination of data augmentation, Bi-atten-ResUnet, and post-processing to achieve the image segmentation task of ovarian tumors. The data augmentation method based on rotation, translation, etc. expands the capacity of the training dataset and reduces the risk of model overfitting. Then, to solve the problem of limited number of samples and the complicated background, Bi-atten-ResUnet is proposed, and it is one of the innovative points of this thesis. Based on the U-net network, Bi-atten-ResUnet uses transfer knowledge obtained from the pre-trained ResNet to make the segmentation model easier to train. The usage of bi-non-local attention modules lets the network focus on the effective information of the input image, which improves the segmentation results. The adoption of the post-processing operation reduces the influence of small connected domain noise and solves the problem of misdetection of the bladder. Results show that the proposed model can achieve automatic segmentation of ovarian cancer on MR images.

4 Category Prediction of Ovarian Cancers

4.1 Introduction

After the segmentation of ovarian tumors in MR images is achieved, the next work is to study the benign and malignant classification, Type I and Type II subtype classification and other secondary tasks, such as survival prediction and KI-67 prediction. This chapter proposes a multi-sequence MRI ovarian cancer diagnosis system, which is divided into three steps: quantitative description feature extraction, feature selection, and establishment of prediction models. In terms of feature extraction, the existing feature systems are most dependent on some qualitative features, which leads to too many subjective factors of different radiologists during the diagnosis process. This chapter presents a quantitative description system for ovarian cancer tumors, including radiomic features and network features. Among them, the network features are derived from the segmentation network, which is also one of the innovations in this thesis. The information of the segmentation network is not only used to acquire the lesion area, but also assists the classification tasks, thereby improving the utilization efficiency of the feature.

Feature selection operation is necessary because of the large number of features acquired, and the intricate relationships between features, such as redundancy, correlation, and the existence of the dimension disaster problem. In this chapter, the MR image features of ovarian cancer are selected via iterative sparse representation (ISR) method to reduce the feature vector dimension. The selected effective features are used to establish a predictive model to complete diagnostic tasks for ovarian cancer.

4.2 Evaluation Methods

Due to the different targets, a variety of metrics are needed to fully evaluate the models from multiple aspects. The benign and malignant tumor diagnosis and Type I and Type II OEC categorization are binary classification problems. The Ki-67 prediction is converted to the binary classification tasks by setting the threshold 50. In this thesis, we apply area under the receiver operating characteristic curve (AUC), accuracy, sensitivity (SENS) and specificity (SPEC) to evaluate these tasks. Their definitions are as follows.

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{4-1}$$

$$SENS = \frac{TP}{TP + FN} \tag{4-2}$$

$$SPEC = \frac{TN}{TN + FP} \tag{4-3}$$

TP denotes true positive, TN denotes true negative, FP symbolizes false positive, and FN symbolizes false negative.

As for AUC calculation, the area enclosed by the true positive rate (TPR) and false positive rate (FPR) axes is called AUC. This method is not affected by sample imbalance problem and therefore it is a stable metric.

In survival analysis tasks, patients are categorized into two groups: low risk and high risk. Hazard ratio and Kaplan-Mier (KM) are applied to evaluate the model performance. The larger the value of hazard ratio, the greater the difference in prognosis between the two groups, and the better the classification performance of the model. By observing the KM plots, we can intuitively understand the survival conditions of two groups of patients.

4.3 Feature Extraction

4.3.1 Radiomic Features

Radiomic features are extracted from lesions of the ovarian image according to

segmentation result provided in the Chapter Three. The obtained features mainly fall into 4 groups: intensity, shape, texture and wavelets.

(1) The intensity group consists of 22 features, describing the overall statistical intensity information in the segmented lesion area.

(2) The shape group contains 15 features, obtaining the structure of the boundary of the lesion ROI [36].

(3) The texture group comprises 39 features, describing the gray-level regional spatial distribution. Comparing to the intensity features targeted to the overall information, texture features focused more on locality, and they not only acquire the frequency but also the location of pixel [37].

(4) In the wavelet group, intensity group and texture group features are transformed into eight frequency sub-bands via wavelet to acquire some hidden characteristics [38]. Altogether, 488 wavelet features are extracted by this operation. The details of high-throughput features are summarized in Table 4-1.

For axial T1WI, sagittal fs-T2WI and coronal T2WI MRI images, all four groups of features are obtained. Meanwhile, only intensity features extraction is applied in ADC map, since lesions in ADC map are blurred and distorted, only the intensity information is useful for analysis. Totally, we extracted 1714 features for each lesion from 4 sequences of MR images.

Table 4-1 Summary of 564 radiomics features

Feature category			**Feature numbers**
Intensity Group			**22**
1) energy	*2) entropy*	*3)kurtosis*	*4) max*
5) mean deviation	*6) mean*	*7) media*	*8) min*
9) range	*10) root mean square*	*11) skewness*	*12) standard deviation*
13) h-uniformity	*14) h-variance*	*15) h-mean*	*16) h-variance*
17) h-skewness	*18) h-kurtosis*	*19) 10th percentile*	*20) 25th percentile*
21) 75th percentile	*22) 90th percentile*		
Shape Group			**15**
1) compactness	*2) compacted square*	*3) max length*	*4) disproportion*
5) sphericity	*6) superficial area*	*7) surface to volume ratio*	*8) volume*
9) region to bounding	*10) major length*	*11) minor length*	*12) eccentricity*
13) orientation	*14) solidity*	*15) Fourier descriptors*	
Texture Group			**9**
GLCM(gray-level co-occurrence matrix)			
1) energy	*2) contrast*	*3) entropy*	*4) homogeneity*
5) correlation	*6) sum average*	*7) variance*	*8) dissimilarity*
GLRLM(gray-level run-length matrix)			
1) run-length variance	*2) long run emphasis*	*3) run percentage*	*4) length-uniformity*
5) gray-level nonuniformity	*6) low gray-level run emphasis*	*7) high gray-level run emphasis*	*8) short run low gray-level emphasis*
9) short run high gray-level emphasis	*10) long run low gray level emphasis*	*11) long run high gray-level emphasis*	*12) gray-level variance*
13) short run emphasis			
GLSZM(gray-level size zone matrix)			
1) small zone emphasis	*2) large zone emphasis*	*3) gray-level nonuniformity*	*4) zone-size nonuniformity*
5) zone percentage	*6) low gray-level zone emphasis*	*7) high gray-level zone emphasis*	*8) small zone low gray-level emphasis*
9) small zone high gray-level emphasis	*10) large zone low gray-level emphasis*	*11) large zone high gray-level emphasis*	*12) gray-level variance*
13) zone-size variance			
NGTDM(neighborhood gray-tone difference matrix)			
1) coarseness	*2) contrast*	*3) busyness*	*4) complexity*
5) strength			
Wavelet Group			**48**
LLL HLL LHL HHL LLH HLH LHH HHH decomposition for intensity and texture features			
Total feature number			**564**

4.3.2 Network Features

The network feature is actually the parameter of the middle layer of the deep neural network. These high dimensional features are usually difficult to process directly. However, after the dimensionality reduction, these features can be treated as normally obtained features, and processed by machine learning algorithms, such as classifiers. As shown in Figure 4-1, the network features are captured from the exactly middle layer of Bi-atten-ResUnet. The dimension reduction method is global average pooling (GAP) in this thesis [39].

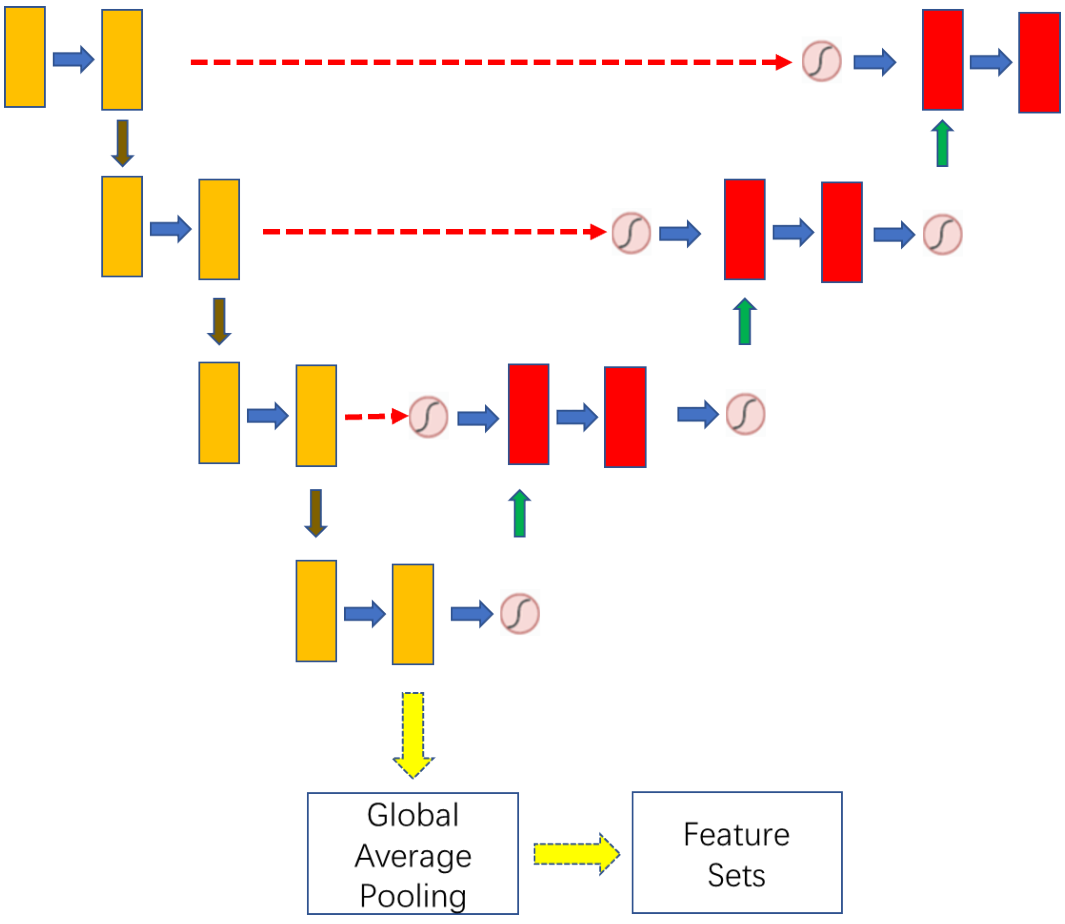

Figure 4-1 Network Features

In deep learning tasks, GAP is mainly used to reduce the number of neurons in the last several layers of the network to solve the overfitting problem caused by too many

parameters in the fully connected layer. GAP calculates the average value of each 2-dimensional feature map, and applies the value to represent the information of the entire map. In this way, the original high-dimensional $h*w*d$ feature map is compressed to $1*1*d$. Specifically, in this experiment, the dimension of feature map before GAP is $224*224*16$, and output of GAP is 16, which drastically reduces the number of descriptors to facilitate further processing. These 16 network features will be combined with the radiomic features to establish the diagnostic model.

4.4 Feature Selection

Some features are highly correlated with others and high dimensional features increase the computational complexity in the classification task. Also, due to the complexity of the feature space, there exists spatial overlapping of features, making it difficult to distinguish different types of samples in space [40], [41]. Therefore, the feature selection operation for original features is required. In order to obtain the most effective features subset, we apply the ISR feature selection method for all extracting features [42]. The sparse representation (SR) method intends to obtain a group of features relevant to the target rather than taking the performance of each feature into account alone [43]. Moreover, ISR enhances efficiency and robustness of feature selection by randomly picking a part of samples to do the filtering operation ahead during each iteration. The sparse representation coefficient (SRC) which denotes the importance of the corresponding feature is calculated. Particularly, features with zero coefficient will be abandoned for building models. The calculation process of ISR is as follows.

The formula for calculating the coefficient can be described as follows:

$$\hat{\alpha} = \arg\min_{\alpha} \|l - F\alpha\|_2^2 + \eta \|\alpha\|_1 \tag{4-4}$$

Where $\hat{\alpha}$ denotes the coefficient of each feature; $l \in R^m$ represents the true label; m is the size of samples. $F \in R^{m \times n}$ indicates the feature group; n is the number of features. η denotes the regularization parameter. If the feature set has been

normalized, the absolute value of the coefficient calculated symbolizes the importance of the corresponding feature. Features with low-ranked coefficient is considered redundant features and are tended to be removed.

According to the traditional method, SRC is obtained by one iterative operation. It doesn't reflect the importance of characteristic, when the number of samples are far less than the number of features. The ISR method selects a part of features to conduct the sparse representation for each iteration.

$$\hat{\alpha}^{(k)} = \arg\min_{\alpha} \left\| l^{(k)} - F^{(k)}\alpha \right\|_2^2 + \eta \left\| \alpha \right\|_1 \tag{4-5}$$

Where $l^{(k)}$, $F^{(k)}$ and $\hat{\alpha}^{(k)}$ indicate the selected label, feature set and the sparse coefficient. Then we calculate the average result of the coefficient for k iterations.

$$\alpha^{(k)} = \frac{1}{k}\sum \hat{\alpha}^{(k)} \tag{4-6}$$

When the distance between adjacent results is smaller than the predefined threshold $\left\| \alpha^{(k)} - \alpha^{(k-1)} \right\|_2 < \varepsilon$, the iteration stops and we get the final $\alpha^{(k)}$ to select features.

4.5 Classifier

The selected features described in previous paragraph are the final feature set. These features are used to build models for classification tasks, including benign and malignant classification, Type I and Type II subtype classification, KI-67 prediction and survival prediction. The machine learning classifier chosen in this article is support vector machine (SVM).

SVM has been widely used in machine learning. SVM enjoys high discrimination and stability in searching for optimal hyperplane in high dimensional space [44]. SVM is a binary classification model, and its basic model is a linear classifier to search for the maximal interval in the feature space to separate the two categories. By applying a kernel function, SVM can be converted to a nonlinear classifier [45]. The target function of SVM can be written as follows:

$$\begin{aligned} &\min \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{l}\xi \\ &\text{s.t. } y_i(\omega x_i + b) \geq 1 - \xi_i, \quad i = 1,2,...,l \\ &\quad \xi_i > 0, \quad i = 1,2,...,l \end{aligned} \tag{4-7}$$

Where ω is the weights vector, b is the bias vector, ξ is the loosen factor and C is the penalty coefficient. x_i is the i_{th} sample and y_i symbolizes the corresponding true label.

SVM is generally used to balanced data sets. However, for machine learning tasks, if the number of samples in one category in the training set is much larger than the number of another category, the situation is called imbalanced sample distribution and that will cause certain problems. In extreme cases, even if all samples are identified as majority labels, the overall accuracy will not be particularly low. Therefore, for imbalanced samples, special solutions are needed. In this thesis, Cost-Support Vector Classification (C-SVC) algorithm is used to adaptively adjust the penalty factor which is according to the proportion of different categories [46]. The target function of C-SVC can be written as follows:

$$\min \frac{1}{2}\|\omega\|^2 + C_+ \sum_{y_i=1} \xi_i + C_- \sum_{y_i=-1} \xi_i \tag{4-8}$$

Where C_+ and C_- represent respectively the penalty parameters corresponding to the majority and minority samples. With the adoption of C-SVM, the final diagnostic model becomes more efficient and more generalizable. The experiment is divided into two parts. The leave-one-out cross-validation (LOOCV) is first applied to verify and adjust the model. Then, the independent testing set is used to further evaluate the diagnostic performance of the final model.

4.6 Result

In this thesis, we apply the model to several ovarian diagnostic tasks: benign and malignant classification, Type I and Type II OEC classification, Ki-67 classification with threshold 50 as well as the survival prediction. The results of the experiment are

as follows.

Table 4-2 illustrates the final classification results by the LOOCV and the independent testing groups for three classification tasks. The model achieves an accuracy of 0.906 and an AUC of 0.967 of in distinguishing between benign and malignant ovarian tumors. In Type I and Type II OEC differentiation task, the proposed model achieves an accuracy of 82% and an AUC of 83%. When 50 is used as the cut-off threshold of Ki-67 expression, an accuracy of 75% and an AUC of 74% are achieved in the independent testing cohort. The ROC curves for three classification tasks in the independent testing set are demonstrated in Figure 4-2.

Table 4-2. The diagnostic performance of all selected MRI radiomics features for various classification tasks.

	Classification	AUC	ACC	SENS	SPEC
LOO cross-validation	Benign and Malignancy	0.975	0.903	0.944	0.789
	Type I and Type II	0.859	0.927	0.905	0.950
	KI-67 (threshold 50)	0.806	0.788	0.883	0.553
Independent testing	Benign and Malignancy	0.967	0.906	0.903	0.913
	Type I and Type II	0.823	0.833	0.765	0. 865
	KI-67 (threshold 50)	0.738	0.754	0.840	0.500

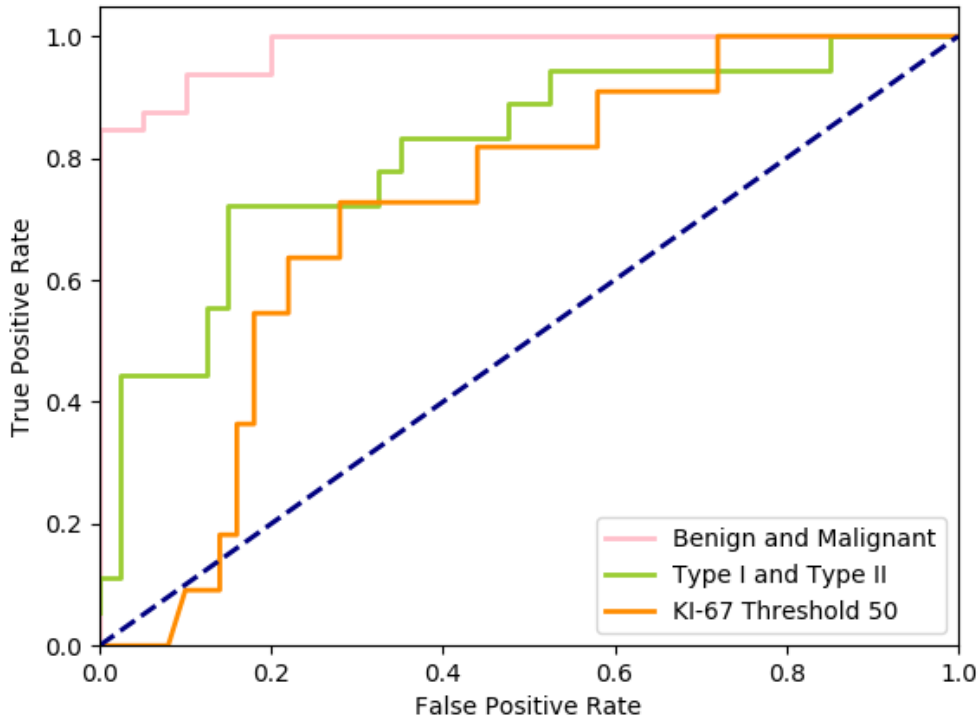

Figure 4-2 Receiver operating characteristic (ROC) curves of the classification tasks

Table 4-3 shows the proposed model in this thesis performs better than the radiologists do in distinguishing malignant tumors from benign tumors. Our model achieves an accuracy of 90.6%, which is higher than the result of experienced radiologists with 83.5%. The most common faults occurred clinically is to mis-classify borderline ovarian tumors (BOTs) into the benign group, which may cause medical risks. In the condition of the exclusion of BOTs, the proposed diagnostic model still shows the better performance than radiologists. The result illustrates the model's potential value clinically.

Table 4-3. Diagnostic performance comparison between radiologist and computer in determining malignant ovarian lesions

	TP	TN	FP	FN	SEN	SPE	ACC
Reader's diagnostic performance (include BOT)	51	20	3	11	0.823	0.869	0.835
Computer with all selected sequence	56	21	2	6	0.903	0.913	0.906

Reader's diagnostic performance (not include BOT)	42	20	3	2	0.955	0.869	0.925
Computer with all selected sequence	42	21	2	2	0.955	0.913	0.940

.

To observe the individual performance of each MRI sequence, we take the same approach on each single sequence. When we only use the radiomic features extracted from a single sequence to build the model, the result is obviously inferior to that of multi-sequence models. The results show the superiority of the multi-sequence approach over the single-sequence approach. Moreover, the model with the combination of radiomic features and deep features achieves the best performance. The results are shown in Table 4-4.

Table 4-4. Diagnostic performance of MRI radiomics in each modality.

Task	Metric	Axi-T1WI	Cor-T2WI	Sag-fs-T2WI	ADC	Multi-sequence	Multi-sequence + deep features
Benign and Malignant	AUC	0.914	0.902	0.860	0.765	0.945	0.967
	ACC	0.848	0.816	0.801	0.766	0.871	0.905
Type I and Type II	AUC	0.754	0.757	0.819	0.730	0.815	0.823
	ACC	0.712	0.712	0.775	0.702	0.802	0.833
KI-67 (threshold 50)	AUC	0.717	0.709	0.633	0.610	0.721	0.738
	ACC	0.703	0.653	0.616	0.557	0.733	0.754

The survival predictions given by our model are shown in Figure 4-3 KM plots. The model divides patients into good and poor prognosis groups. Patients categorized into high risk groups are more likely to suffer from disease progression (HR = 4.362, p = 0.0013).

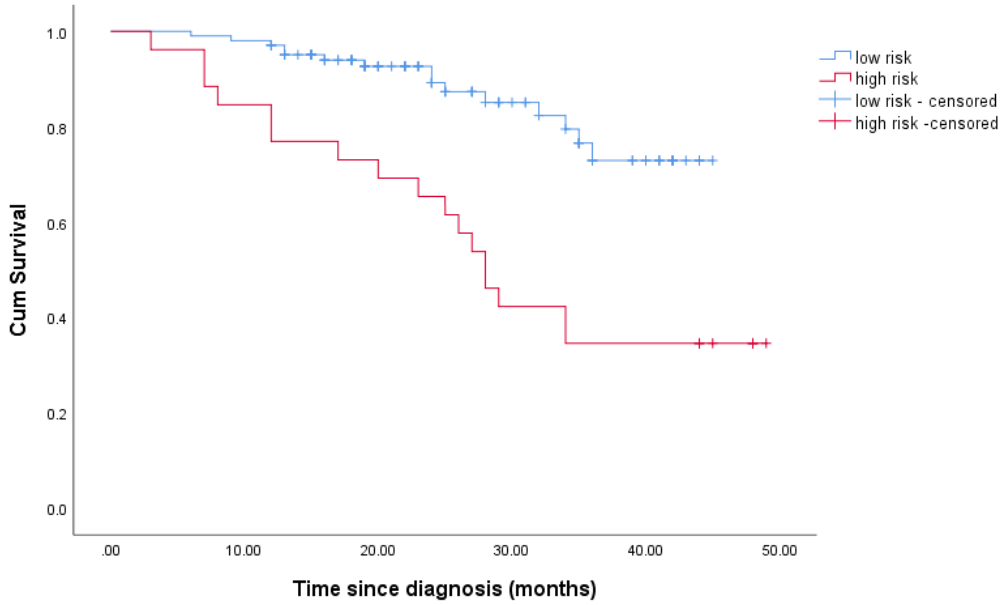

Figure 4-3 Survival Analysis Result

The results show that the proposed diagnostic model achieves an excellent performance in classifying benign and malignant ovarian tumors and in differentiating Type I and Type II OEC in the independent testing set. In addition, the ovarian MRI features demonstrated an excellent ability to categorize patients into high-risk and low-risk groups.

4.7 Summary

In this chapter, a diagnostic model based on the tumor ROI is proposed to accomplish the tasks related to ovarian cancers. The model can be divided into three steps. First, four categories of radiomic characteristics: intensity, shape, texture and wavelet features are extracted from the lesion area. Also, we apply the GAP method to the parameters of the middle layers of the segmentation network, and obtain network features. Then, all these features are selected by ISR method and the chosen features are used for classification by SVM. The model achieves good results in the classification of benign and malignant ovarian cancer, the classification of Type I and Type II OEC, ki-67 prediction, and the survival prediction. Our results show that the

multi-sequence MR model is superior to the single-sequence model. Moreover, the addition of network features can improve the classification performance. We also use the ROC diagram and KM plots to demonstrate the prediction results of the model intuitively.

5 Conclusion and Future Work

5.1 Conclusion

This thesis proposes an automatic diagnosis system of ovarian cancer based on the combination of deep learning and radiomics. The lesion segmentation model for ovarian lesions is based on deep learning segmentation and the prediction of the category of ovarian tumors is based on radiomic approach.

Bi-atten-ResUnet is proposed in this thesis as the segmentation model. The network is established on the basis of U-Net with adopting Residual block and non-local attention module. The residual block helps to solve the vanishing gradient problem, while the attention module focuses the model's attention on the tumor area. The usage of these techniques with the assistance of the data augmentation and post-processing operations lets the model achieve 0.918 Dice coefficient in segmenting the ovarian tumors.

After the segmentation work, the thesis proposes the ovarian diagnostic model based on the ROI segmentation result. This model consists of three steps: feature extraction, feature selection, and establishment of prediction models. First, radiomic features, network features are obtained. Then ISR method is used to eliminate the unrelated and redundant features. Selected effective features are used to establish a predictive model to complete several diagnostic tasks for ovarian cancer. The model achieves an accuracy of 0.906 and an AUC of 0.967 in distinguishing between benign and malignant ovarian tumors, an accuracy of 0.833 and an AUC of 0.823 in distinguishing between Type I and Type II OEC, an accuracy of 75%, an AUC of 74% in KI-67 classification and a hazard ratio of 4.362 between high-risk and low-risk groups.

5.2 Future Work

The ovarian cancer diagnostic model proposed in this thesis still has room for improvement in the future. The model can be improved from the following perspectives.

(1) From the perspective of data sources, due to the limited capability of MR scanners, we have not acquired some MR sequences with latest technology, such as CE-MRI, which is effective in tumor diagnosis [47] For the same reason, all MR images were acquired in a 1.5-T MRI scanner, while 3.0-T MRI machines producing higher resolution images are widely used in other clinical centers. In the future, the model based on the higher-definition sample data may produce better performance.

(2) The non-local attention module in the segmentation model is used to improve the accuracy of lesion segmentation. However, it also brings the heavy computation and the increased processing time for each slice. In the future, when the model is actually deployed on the business hospital platforms, these problems may hinder its further widespread application. Therefore, to find a smaller and more effective attention mechanism will be an important direction for the algorithm improvement. This requires to reconstruct the network architecture and design a new network module.

References

1. Chen, W.Q., et al., *Cancer Statistics in China, 2015.* Ca-a Cancer Journal for Clinicians, 2016. **66**(2): p. 115-132.
2. Park, S.B., et al., *Ovarian serous surface papillary borderline tumor: characteristic imaging features with clinicopathological correlation.* British Journal of Radiology, 2018. **91**(1088).
3. Zhang, H., et al., *Magnetic resonance imaging radiomics in categorizing ovarian masses and predicting clinical outcome: a preliminary study.* European Radiology, 2019. **29**(7): p. 3358-3371.
4. Chen, X.X., et al., *CA-125 Level as a Prognostic Indicator in Type I and Type II Epithelial Ovarian Cancer.* International Journal of Gynecological Cancer, 2013. **23**(5): p. 815-822.
5. Menon, U., M. Griffin, and A. Gentry-Maharaj, *Ovarian cancer screening-Current status, future directions.* Gynecologic Oncology, 2014. **132**(2): p. 490-495.
6. Earle, C.C., et al., *Effect of surgeon specialty on processes of care and outcomes for ovarian cancer patients.* Journal of the National Cancer Institute, 2006. **98**(3): p. 172-180.
7. Elattar, A., et al., *Optimal primary surgical treatment for advanced epithelial ovarian cancer.* Cochrane Database of Systematic Reviews, 2011(8).
8. Iyer, V.R. and S.I. Lee, *MRI, CT, and PET/CT for Ovarian Cancer Detection and Adnexal Lesion Characterization.* American Journal of Roentgenology, 2010. **194**(2): p. 311-321.
9. Wang, W., et al., *Ovarian mucinous borderline tumor accompanied by LGESS with myxoid change: a case report and literature review.* European Journal of Medical Research, 2017. **22**.
10. Aldrich, J.E., *Basic physics of ultrasound imaging.* Critical Care Medicine, 2007. **35**(5): p. S131-S137.
11. Medeiros, L.R., et al., *Accuracy of magnetic resonance imaging in ovarian tumor: a systematic quantitative review.* American Journal of Obstetrics and Gynecology, 2011. **204**(1).
12. Suh, D.H., et al., *Circulating tumor cells in the differential diagnosis of adnexal masses.* Oncotarget, 2017. **8**(44): p. 77195-77206.
13. Gillies, R.J., P.E. Kinahan, and H. Hricak, *Radiomics: Images Are More than Pictures, They Are Data.* Radiology, 2016. **278**(2): p. 563-577.
14. Li, X.Z., et al., *Deep neural network for plasmonic sensor modeling.* Optical Materials Express, 2019. **9**(9): p. 3857-3862.
15. Bourguet, J.R., et al., *An artificial intelligence-based approach to deal with argumentation applied to food quality in a public health policy.* Expert Systems with Applications, 2013. **40**(11): p. 4539-4546.
16. Krizhevsky, A., I. Sutskever, and G.E. Hinton, *ImageNet Classification with Deep Convolutional Neural Networks.* Communications of the Acm, 2017. **60**(6): p. 84-90.
17. Litjens, G., et al., *A survey on deep learning in medical image analysis.* Medical Image Analysis, 2017. **42**: p. 60-88.
18. Qiu, Y., et al. *An initial investigation on developing a new method to predict short-term breast cancer risk based on deep learning technology*. in *Medical Imaging 2016: Computer-Aided Diagnosis*. 2016. International Society for Optics and Photonics.
19. Chlebus, G., et al., *Automatic liver tumor segmentation in CT with fully convolutional neural*

networks and object-based postprocessing. Scientific Reports, 2018. **8**.

20. Bardou, D., K. Zhang, and S.M. Ahmad, *Classification of Breast Cancer Based on Histology Images Using Convolutional Neural Networks.* Ieee Access, 2018. **6**: p. 24680-24693.
21. Rafiei, M., T. Niknam, and M.H. Khooban, *Probabilistic Forecasting of Hourly Electricity Price by Generalization of ELM for Usage in Improved Wavelet Neural Network.* Ieee Transactions on Industrial Informatics, 2017. **13**(1): p. 71-79.
22. Bourgier, C., et al., *Radiomics: Definition and clinical development.* Cancer Radiotherapie, 2015. **19**(6-7): p. 532-537.
23. Lambin, P., et al., *Radiomics: Extracting more information from medical images using advanced feature analysis.* European Journal of Cancer, 2012. **48**(4): p. 441-446.
24. Kazerooni, A.F., et al., *Semiquantitative Dynamic Contrast-Enhanced MRI for Accurate Classification of Complex Adnexal Masses.* Journal of Magnetic Resonance Imaging, 2017. **45**(2): p. 418-427.
25. Rizzo, S., et al., *Radiomics of high-grade serous ovarian cancer: association between quantitative CT features, residual tumour and disease progression within 12 months.* European Radiology, 2018. **28**(11): p. 4849-4859.
26. Qiu, Y.C., et al., *Early prediction of clinical benefit of treating ovarian cancer using quantitative CT image feature analysis.* Acta Radiologica, 2016. **57**(9): p. 1149-1155.
27. Bi, W.L., et al., *Artificial intelligence in cancer imaging: Clinical challenges and applications.* Ca-a Cancer Journal for Clinicians, 2019. **69**(2): p. 127-157.
28. Ronneberger, O., P. Fischer, and T. Brox. *U-net: Convolutional networks for biomedical image segmentation*. in *International Conference on Medical image computing and computer-assisted intervention*. 2015. Springer.
29. Sun, W.W. and R.S. Wang, *Fully Convolutional Networks for Semantic Segmentation of Very High Resolution Remotely Sensed Images Combined With DSM.* Ieee Geoscience and Remote Sensing Letters, 2018. **15**(3): p. 474-478.
30. He, K., et al. *Deep residual learning for image recognition*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
31. He, K., et al. *Identity mappings in deep residual networks*. in *European conference on computer vision*. 2016. Springer.
32. Veit, A., M. Wilber, and S. Belongie, *Residual Networks Behave Like Ensembles of Relatively Shallow Networks.* Advances in Neural Information Processing Systems 29 (Nips 2016), 2016. **29**.
33. Wu, Z.X., et al., *BlockDrop: Dynamic Inference Paths in Residual Networks.* 2018 Ieee/Cvf Conference on Computer Vision and Pattern Recognition (Cvpr), 2018: p. 8817-8826.
34. Bahdanau, D., K. Cho, and Y. Bengio, *Neural machine translation by jointly learning to align and translate.* arXiv preprint arXiv:1409.0473, 2014.
35. Wang, X., et al. *Non-local neural networks*. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018.
36. Hu, Y.Z., et al., *Reproducibility of quantitative high-throughput BI-RADS features extracted from ultrasound images of breast cancer.* Medical Physics, 2017. **44**(7): p. 3676-3685.
37. Ortiz-Ramon, R., et al., *Classifying brain metastases by their primary site of origin using a radiomics approach based on texture analysis: a feasibility study.* European Radiology, 2018. **28**(11): p. 4514-4523.

38. Aerts, H.J.W.L., et al., *Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach (vol 5, pg 4006, 2014).* Nature Communications, 2014. **5**.
39. Lin, M., Q. Chen, and S. Yan, *Network in network.* arXiv preprint arXiv:1312.4400, 2013.
40. Unler, A., A. Murat, and R.B. Chinnam, *mr(2)PSO: A maximum relevance minimum redundancy feature selection method based on swarm intelligence for support vector machine classification.* Information Sciences, 2011. **181**(20): p. 4625-4641.
41. Sakar, C.O., O. Kursun, and F. Gurgen, *A feature selection method based on kernel canonical correlation analysis and the minimum Redundancy-Maximum Relevance filter method.* Expert Systems with Applications, 2012. **39**(3): p. 3432-3437.
42. Wu, G.Q., et al., *Sparse Representation-Based Radiomics for the Diagnosis of Brain Tumors.* Ieee Transactions on Medical Imaging, 2018. **37**(4): p. 893-905.
43. Li, Y.Q., et al., *Voxel Selection in fMRI Data Analysis Based on Sparse Representation.* Ieee Transactions on Biomedical Engineering, 2009. **56**(10): p. 2439-2451.
44. Liu, T.T., et al., *Comparison of the application of B-mode and strain elastography ultrasound in the estimation of lymph node metastasis of papillary thyroid carcinoma based on a radiomics approach.* International Journal of Computer Assisted Radiology and Surgery, 2018. **13**(10): p. 1617-1627.
45. Agarwal, K., et al., *Process knowledge based multi-class support vector classification (PK-MSVM) approach for surface defects in hot rolling.* Expert Systems with Applications, 2011. **38**(6): p. 7251-7262.
46. Zhang, C.H., X.J. Shao, and D.W. Li, *Knowledge-based Support Vector Classification Based on C-SVC.* First International Conference on Information Technology and Quantitative Management, 2013. **17**: p. 1083-1090.
47. Lin, G., et al., *Comparison of the Diagnostic Accuracy of Contrast-Enhanced MRI and Diffusion-Weighted MRI in the Differentiation Between Uterine Leiomyosarcoma/Smooth Muscle Tumor With Uncertain Malignant Potential and Benign Leiomyoma.* Journal of Magnetic Resonance Imaging, 2016. **43**(2): p. 333-342.

Published Work

1. He Zhang[#], **Yunfei Mao**[#], Xiaojun Chen, Guoqing Wu, Xuefen Liu, Peng Zhang, Yu Bai, Pengcong Lu, Weigen Yao, Yuanyuan Wang, Jinhua Yu, Guofu Zhang. Magnetic resonance imaging radiomics in categorizing ovarian masses and predicting clinical outcome: a preliminary study [J]. European Radiology, 2019, 29(7): 3358-3371. (SCI)
   He Zhang and Yunfei Mao contributed equally to the work.

Acknowledgement

I want to extend the sincere appreciation to my two mentors, Assoc Prof. Tapio Pahikkala in UTU and Prof. Jinhua Yu in FDU. They provided me great help and guidance when I wrote and revised my dissertation. I also want to extend gratitude to my friends in the medical signal processing laboratory. They have always been friendly to me and supported me. Thanks to Assoc Prof. Zhang, who providing me the valuable medical knowledge, and we co-published a journal article. Finally, many thanks to my dearest parents, for 25-year love and care