Project Documentation

Team ID	NM2023TMID18418
Project name	CancerVision: Advanced Breast Cancer Prediction with Deep Learning

CancerVision: Advanced Breast Cancer Prediction with Deep Learning

Abstract

The prevalence of breast cancer continues to grow, affecting about 300,000 females in the United States in 2023. However, there are different levels of severity of breast cancer requiring different treatment strategies, and hence, grading breast cancer has become a vital component of breast cancer diagnosis and treatment planning.

Specifically, the gold-standard Scarff-Bloom-Richardson (SBR) grade has been shown to consistently indicate a patient's response to chemotherapy. Unfortunately, the current method to determine the SBR grade requires removal of some cancer cells from the patient which can lead to stress and discomfort along with costly expenses.

In this paper, we study the efficacy of deep learning for breast cancer grading based on synthetic correlated diffusion (CDIs) imaging, a new magnetic resonance imaging (MRI) modality and found that it achieves better performance on SBR grade prediction compared to those learnt using goldstandard imaging modalities.

Hence, we introduce Cancer-Net BCa-S, a volumetric deep radiomics approach for predicting SBR grade based on volumetric CDIs data. Given the promising results, this proposed method to identify the severity of the cancer would allow for better treatment decisions without the need for a biopsy.

Introduction

The prevalence of breast cancer continues to grow, affecting about 300,000 females in the United States in 2023 [1]. However, not all breast cancer is fatal. Specifically, the two main types of breast cancer are in situ and invasive breast cancer [2].

The former is a less severe form of breast cancer that is a precursor to the latter type. The latter type, invasive breast cancer, represents approximately 80% of diagnosed cases and signifies that the cancer can or has already spread into the nearby tissue areas [1, 3].

Patients with invasive breast cancer also often receive a breast cancer grade that represents the similarity of the cancer cells to normal cells under the microscope. The three breast cancer grades (low, intermediate, and high) describe the speed of growth and likelihood of a good prognosis.

Low grade (grade 1) cancer has the best prognosis with slow growth and spread of the cancer, while high grade (grade 3) cancer has the worst prognosis with the greatest difference between cancer and normal cells and represent cancer that is fast-growing with quick spread to other cells.

As such, the stage and grade of breast cancer are vital factors used to determine the severity of breast cancer and discern the best treatment strategy as the stage and grade have been shown to relate to the success of various.

Breast cancer (BC) is the most common cancer affecting women worldwide and the most frequent cause of cancer death in women (1, 2).

Recent advances in treatment strategies have improved BC-related mortality and morbidity; however, almost 30% of BC patients show recurrence in the follow-up.

Therefore, to improve BC outcomes, it is necessary to focus on research such as improving screening methods for early detection of recurrence according to risk stratification, identifying new biomarkers, and developing new innovative treatment strategies.

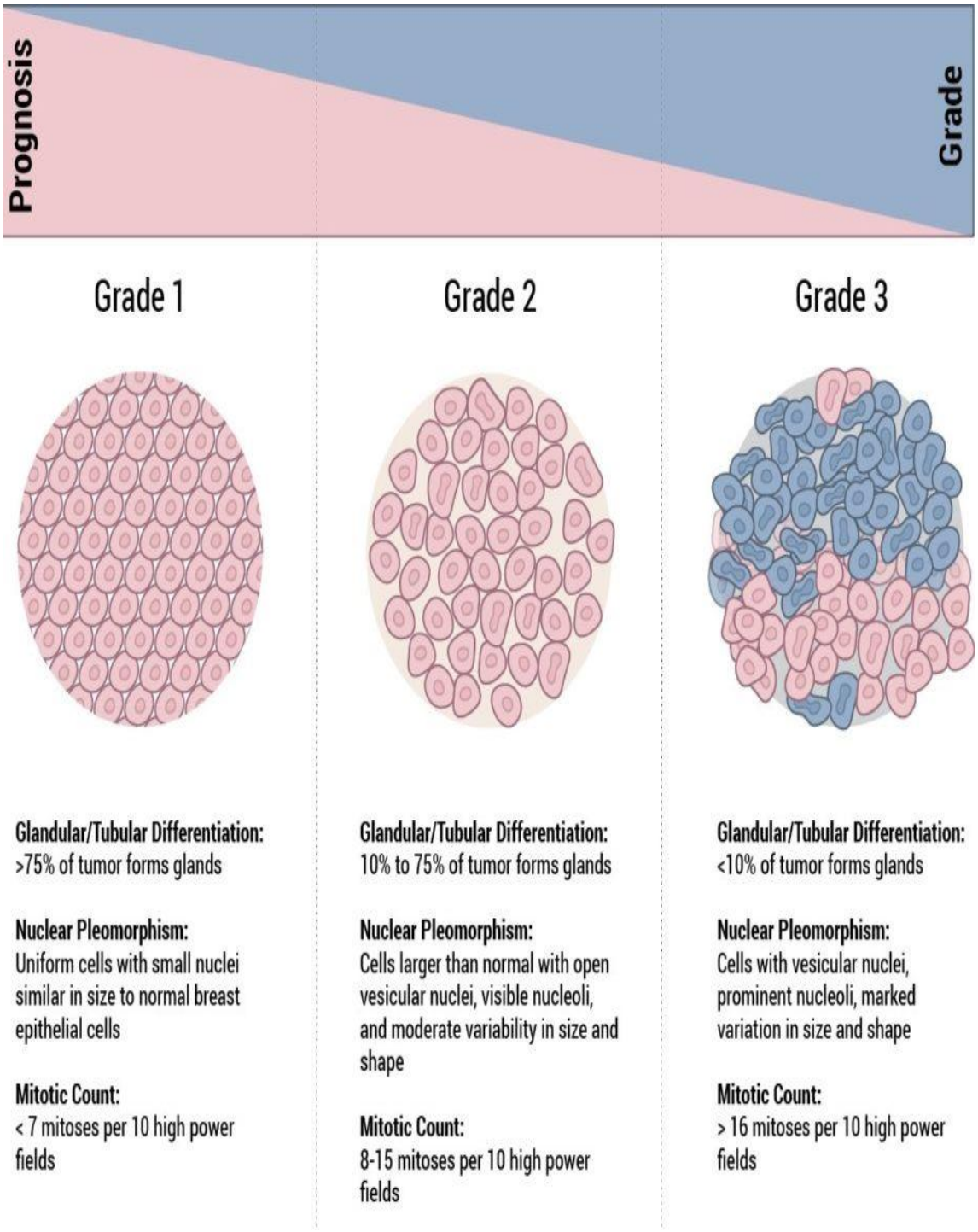


Table 1: SBR grade distribution in the patient cohort

SBR Grade	Number of Patients
Grade I (Low)	5

Grade II (Intermediate)	72
Grade III (High)	175

treatment strategies [4–8]. Specifically, the gold-standard Scarff-Bloom-Richardson (SBR) grade (with example CDIs shown in Fig. 1) has been shown to consistently indicate a patient's response to chemotherapy [9, 10].

Unfortunately, the current method to determine the SBR grade requires removal of some cancer cells from the patient which can lead to stress and discomfort along with costly expenses.

In this paper, we study the efficacy of deep learning for breast cancer grading based on synthetic correlated diffusion (CDIs) imaging, a new magnetic resonance imaging (MRI) modality and found that it achieves better performance on SBR grade prediction compared to those learnt using gold-standard imaging modalities.

Hence, we introduce Cancer-Net BCa-S, a volumetric deep radiomics approach for predicting SBR grade based on volumetric CDIs data. Given the promising results, this proposed method to identify the severity of the cancer would allow for better treatment decisions without the need for a biopsy.

Methodology

The pre-treatment (T0) patient cohort of 252 patient cases across 10 different institutions obtained from the American College of Radiology Imaging Network (ACRIN) 6698/I-SPY2 study [11–14] is used in this study.

More specifically, CDIs acquisitions was attained for each of the patient cases. Furthermore, diffusion-weighted imaging (DWI) acquisitions, T2-weighted (T2w) acquisitions, and apparent diffusion coefficient (ADC) maps were also procured to compare against the performance of CDIs for grade prediction.

The SBR grade from the ACRIN 6698/I-SPY2 study was also obtained for learning and evaluation purposes to compare the current gold-standard MRI modalities and CDIs .

As seen in Table 1, there is an uneven distribution of patients between the three grades and hence, SBR grade I and II were combined into one category.

To investigate the efficacy of volumetric deep radiomic features from CDIs for breast cancer grading, we adapt a previously introduced deep radiomic clinical support workflow [15] to the proposed purpose of predicting the SBR grade of a patient (see Fig. 2).

First, we obtain CDIs acquisitions for each patient by acquiring multiple native diffusion signals with different b-values, passing it into a signal synthesizer to produce synthetic signals and then mixing the native and synthetic signals together to obtain a final signal (CDIs).

To achieve dimensional consistency for machine learning, the CDIs acquisitions are standardized into 224x224x25 volumetric data cubes. We then leverage a 34-layer volumetric residual convolutional neural network architecture trained on breast cancer image volumes [15] to produce deep radiomic features for each patient based on their CDIs data cubes.

A grade predictor composed of a fully-connected neural network architecture is then learnt based on the extracted deep radiomic features from this volumetric network and categorized patient grade (Grade I/Grade II and Grade III).

This grade predictor is used to predict the categorized patient grade prior to treatment.

Leave-one-out cross-validation (LOOCV) is used to evaluate the efficacy of the proposed approach on the patient cohort with accuracy being the main performance metric of interest. The same 34- layer volumetric residual convolutional neural network architecture for volumetric deep radiomic feature extraction .

Table 2: SBR grade prediction accuracy using LOOCV for different imaging modalities

Modality	Accuracy	Sensitivity	Specificity
CDIs	87.70%	90.29%	81.82%
T2w	76.59%	99.43%	24.68%
ADC	69.44%	100.00%	0.00%
DWI	69.44%	95.43%	10.39%

As seen in Table 2, leveraging volumetric deep radiomic features for CDIs achieves the highest grade predictive accuracy of 87.7% with both sensitivity and specificity values over 80%.

Furthermore, CDIs outperforms the gold-standard imaging modalities with an improvement of over 10% on the next highest modality (T2w).

In this paper, we introduce Cancer-Net BCa-S, a volumetric deep radiomics approach for predicting SBR grade based on volumetric CDIs data with a categorized grade prediction accuracy of 87.70%.

With the highest gold-standard MRI modality only achieving a prediction accuracy of 76.59%, over 10% lower than CDIs , the proposed approach with CDIs can increase the grade prediction performance compared to gold-standard MRI modalities.

Given the promising results, future work involves expanding the study with a larger patient cohort to further validate our findings and leveraging improved CDIs coefficient optimization to improve prediction performance.

DATASET AND METHOD

In this section, we first introduce the details of the breast cancer dataset chosen for the study and the pre-processing method, and then introduce the proposed fine segmentation model of the lesion region and the important model components.

Dataset description

The research data of the paper comes from the public dataset officially provided in the Camelyon16 competition.

The Camelyon16 dataset [16] consists of a total of 400 whole slide images (WSIs) split into 270 for training and 130 for testing.

Institution	Train cancer	Train normal	Test
Radboud UMC	90	70	80
UMC Utrecht	70	40	50
Total	160	110	130

The ground truth data for the training slides consists of a pathologist's delineation of regions of metastatic cancer on WSIs of sentinel lymph nodes.

The data was provided in two formats: XML files containing vertices of the annotated contours of the locations of cancer metastases and WSI binary masks indicating the location of the cancer metastasis.

Dataset pre-processing :

Recognition dataset for the region of interest:

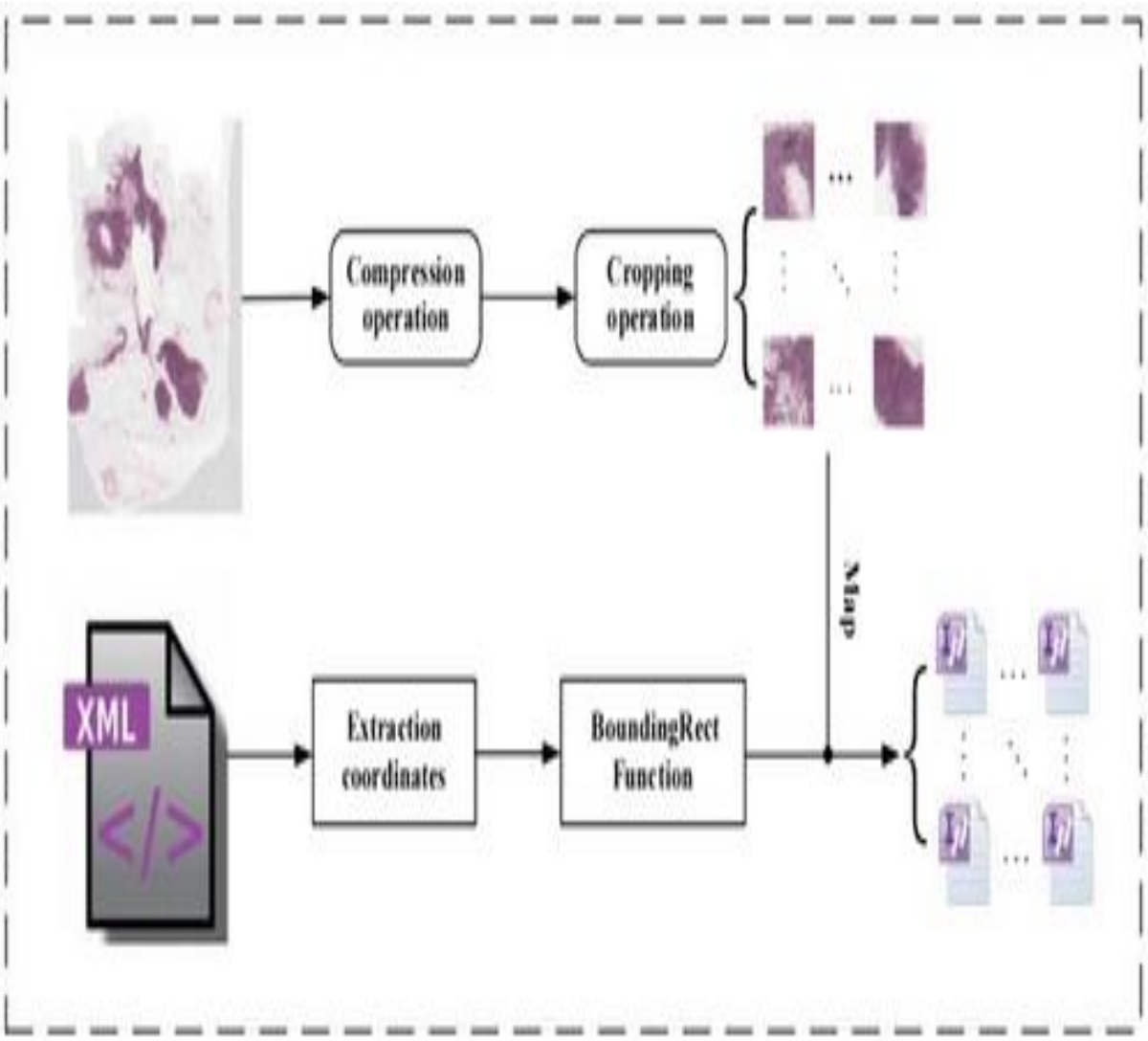
The whole slide images (WSIs) provided by the Camelyon16 dataset are extremely high-resolution, with an average size of 90,000 × 220,000 pixels, which contain a lot of fine-grained information.

We first compress the whole slide images by eight times, extract the edge pixel coordinates of the lesion region in annotation file, and use the boundingRect function to obtain the coordinates of the smallest closed rectangle surrounding the lesion region.

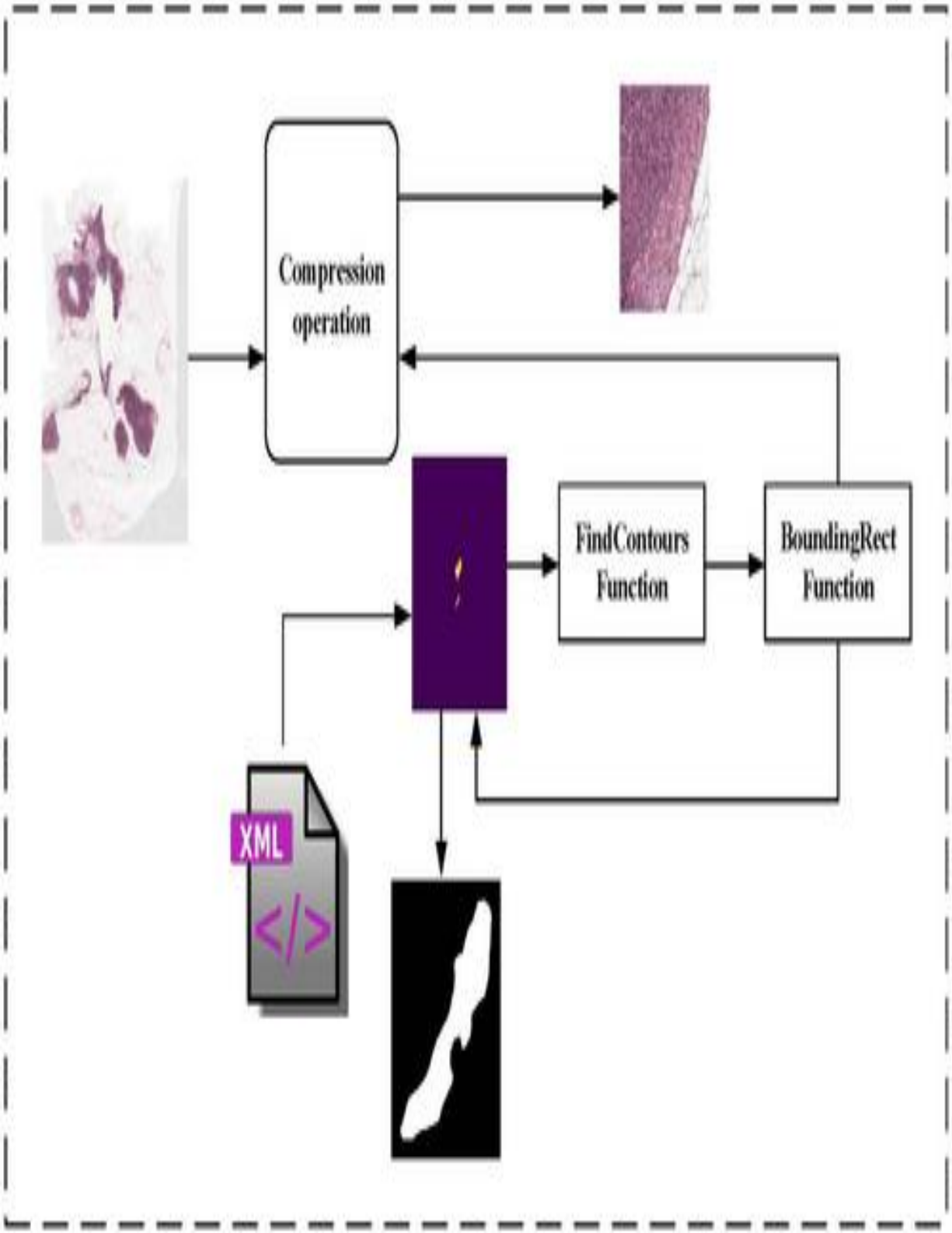
Then, we divide the pathological section into patches by overlapping sliding window, the size of window is 3000 × 3000 pixels, and the size of stride is 1500 × 1500 pixels.

Finally, according to the coordinate mapping relationship between the patch and compressed image, we use coordinates of the smallest closed rectangle to generate the annotation file of patch.

Figure 1 shows the process of generating ROI recognition dataset.



Segmentation dataset for lesion region :



Framework of lesion region segmentation model :

The overall framework of lesion region segmentation model is shown in Figure 3. It consists of three parts involving the image pre-processing module, the ROI recognition module and the lesion region segmentation module.

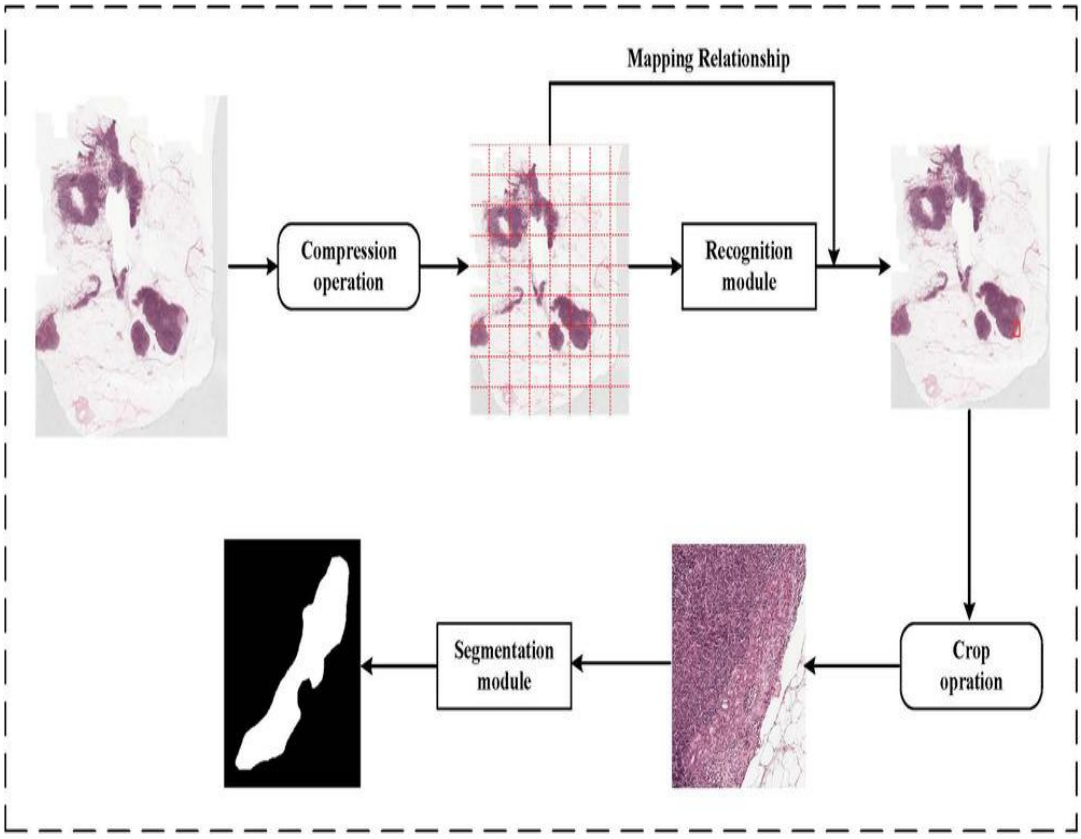
First, we compress the pathological image by eight times and divide it into patches by non-overlapping sliding window, the size of window is 3000 × 3000 pixels.

Then, we use the improved single-stage object detection network YOLO-V4 to recognize the ROI in the processed images.

Based on the mapping relationship between the obtained prediction frame coordinates and the original image, we calculate the corresponding coordinates of ROI in the original image, crop the ROI from the original image and feed the cropped images to the segmentation module.

Finally, we use the global context-aware progressive aggregation network (GCPANet) to do fine segmentation of the lesion region.

The recognition and segmentation module of lesion region will be introduced in detail in the following.



Recognition module for the region of interest :

From R-CNN [18] to Faster R-CNN [19, object detection always follows the idea of region proposal and classification, training two models will lead to an increase in parameters and training cost.

But the regression-based object detection network YOLO does not need to generate a candidate window.

It only needs to send the image into the target window, then the position of the object can be obtained.

YOLO-V4 is a single-stage object detection algorithm, which improves on YOLO-V3 [20], making its speed and accuracy greatly improved.

The main improvements of YOLO-V4 can be divided into three parts. The first part, which uses Mosaic data augmentation operations to improve the training speed and accuracy of the network, and uses cross mini-batch normalization (cmBN) and self-adversarial training (SAT) to improve the generalization performance of the network.

The second part, which uses CSPDarknet53 as the backbone network, and uses the Mish activation function instead of the original RELU activation function, and adds the Dropblock block to further improve the generalization ability of the network.

The third part, which uses the SPP module to fuse feature maps of different scale sizes.

Next, the basic components of YOLO-V4 are briefly introduced, as shown in Figure 3.

The CBL module consists of Conv, BN, Leak_RELU activation function. The CBM is the smallest component of the network structure and consists of Conv, BN, Mish activation function.

The Res unit borrows the residual structure from the ResNet network to build a deep network, and the CBM is a submodule of the residual module.

The CSP_X borrows from the CSPNet network structure, consisting of a convolutional layer and X Res unit modules concatenated. The SPP uses maxpooling of 1×1, 5×5, 9×9 and 13×13 for multi-scale feature fusion.

Intersection over Union (IoU) is a common index for evaluating the accuracy of bounding box localization in object detection tasks.

It reflects the degree of intersection between the predicted box and the ground truth.

The formula is defined as Formula 1, where *B* represents the prediction box, and *G* represents the ground truth.

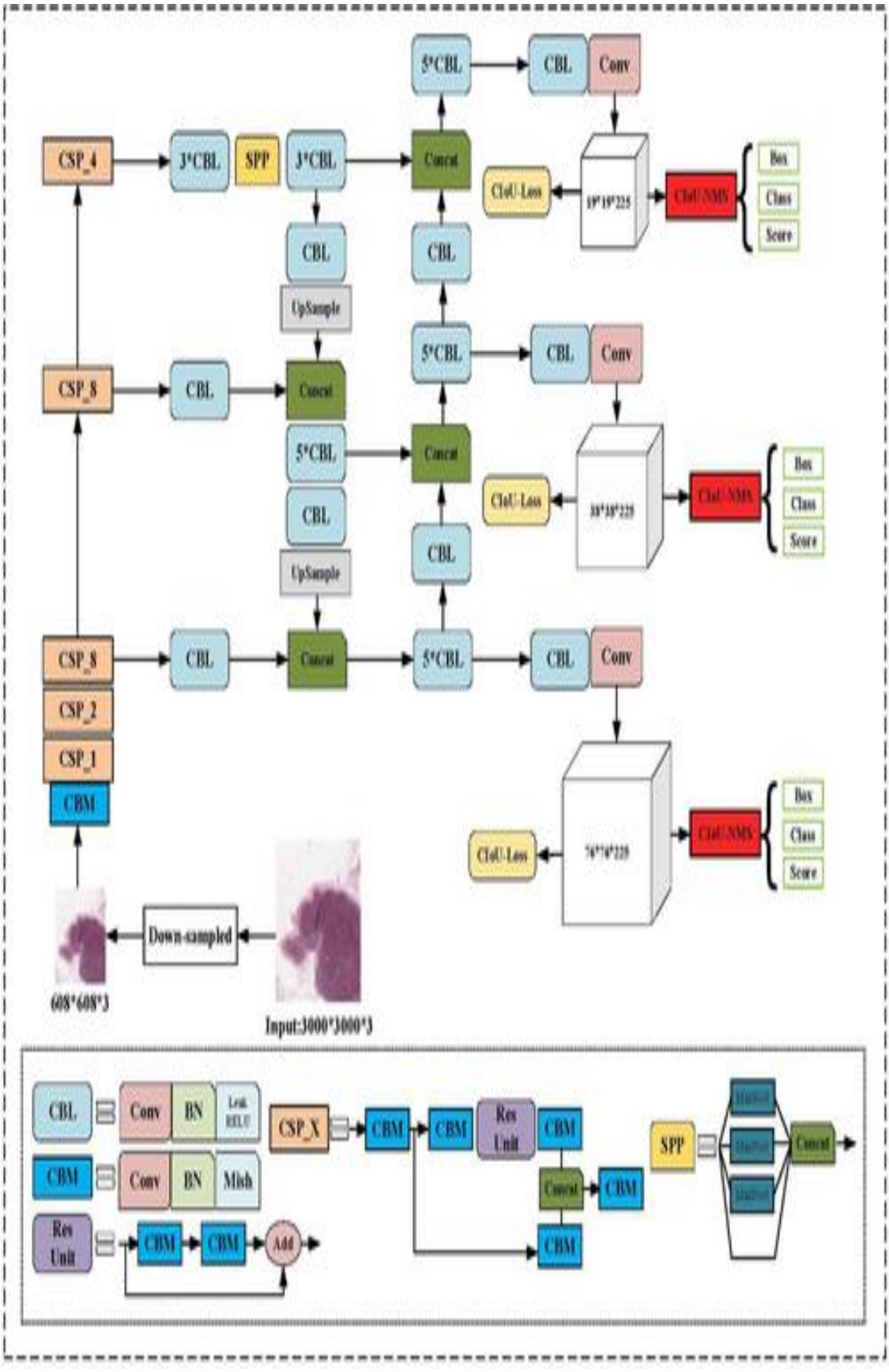
Since there are some problems in using IoU as an evaluation index directly for the loss function, for example, if the predicted frame and the ground truth frame do not intersect, the IoU value will always be equal to zero. To solve the problem, we improve the bounding box regression loss by using CIoU [21], it takes into account geometric factors such as overlapping area, centre point distance, aspect ratio etc.

where represents the Euclidean distance between the centre points of the prediction box and the ground truth, *c* represents the diagonal distance of the smallest closed rectangle that contains both the prediction box and the ground truth, α and *v* represent aspect ratio, *w* and *h* represent the height and width of the prediction box and represent the height and width of the ground truth.

According to the above analysis, we construct the ROI recognition module based on a single-stage object detection network YOLO-V4 and the improvement to bounding box regression loss by using CIoU.

It increases the localization speed and it is also effective for improving the recognition accuracy of ROI.

The specific module composition is shown in the Figure **4**.


5*CBL
CBL
Conv
CSP_4
3*CBL
SPP
3*CBL
Concat
CIoU-Loss
19*19*225
CIoU-NMS
Box
Class
Score
CBL
CBL
UpSample
5*CBL
CBL
Conv
CSP_8
CBL
Concat
5*CBL
Concat
CIoU-Loss
38*38*225
CIoU-NMS
Box
Class
Score
CBL
CBL
UpSample
CSP_8
CBL
Concat
5*CBL
CBL
Conv
CSP_2
CSP_1
CBM
CIoU-Loss
76*76*225
CIoU-NMS
Box
Class
Score
Down-sampled
608*608*3
Input:3000*3000*3
CBL
Conv
BN
Leak. RELU
CBM
Conv
BN
Mish
Res Unit
CBM
CBM
Add
CSP_X
CBM
CBM
Res Unit
CBM
Concat
CBM
CBM
SPP
Concat

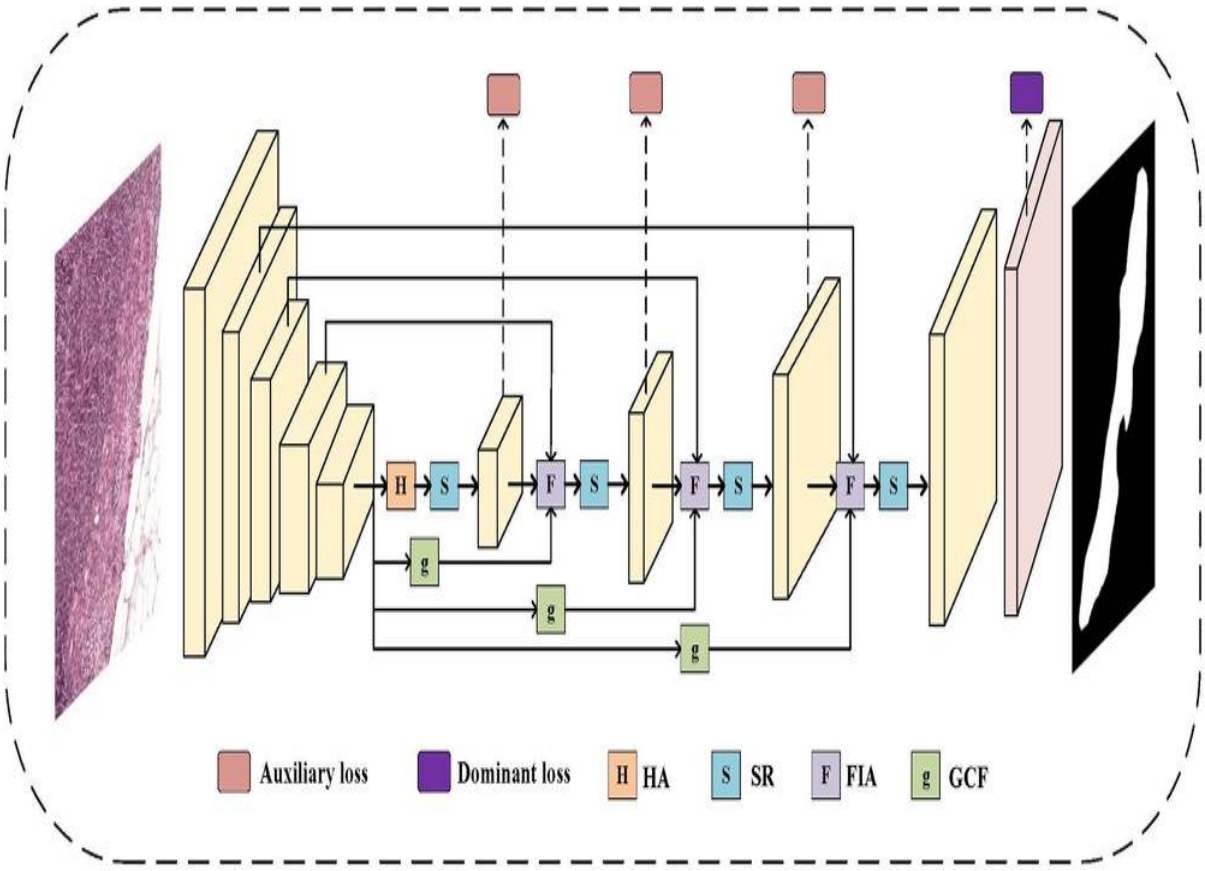
Segmentation module for lesion region :

Most of the previous segmentation network has not effectively utilized the overall context information, which can effectively consider the connection of multiple segmentation regions, resulting in more and more complete segmentation regions.

Considering that lesion region is the largest and most obvious part of the whole image, we construct the lesion region segmentation module based on a Global Context-Aware Progressive Aggregation Network (GCPANet) [22].

It is a symmetrical encoder-decoder architecture, where the encoder component is based on ResNet50 [23] to extract the multi-level features, and the decoder component progressively integrates the multi-level comprehensive features to generate the saliency map in a supervised way.

The specific module composition is shown in the Figure 5.



In saliency detection, binary cross-entropy loss is often used as the loss function to measure the relation between the generated saliency map and the ground truth.

The formula is defined as Formula 5, where *H*, *W* denote the height and width of the image, respectively, is the ground truth label of the pixel)and represents the corresponding probability of being salient objects in position)

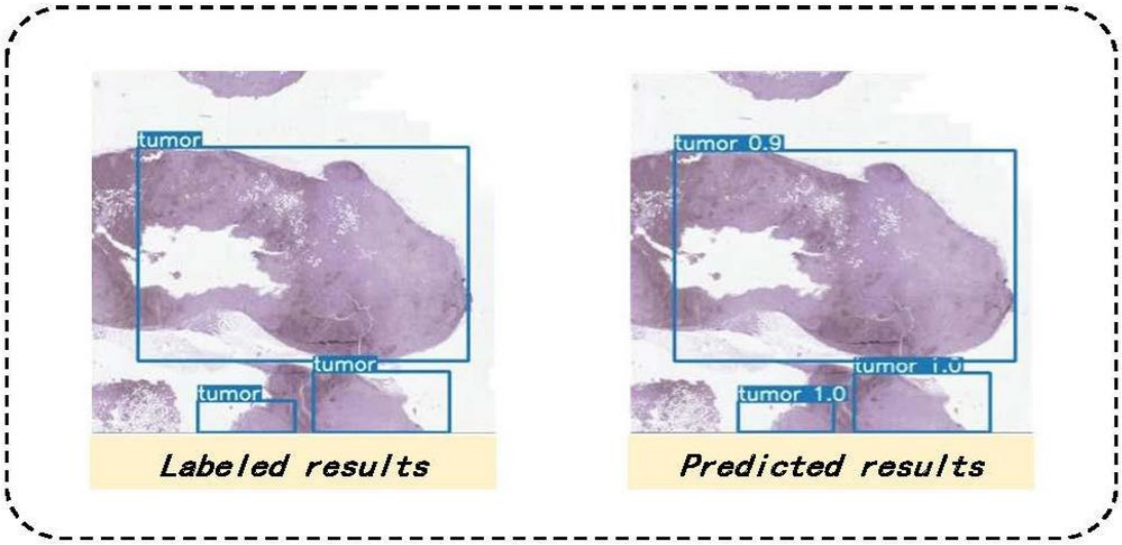
To facilitate the optimization of the segmentation module, we add auxiliary loss at three decoder stages.

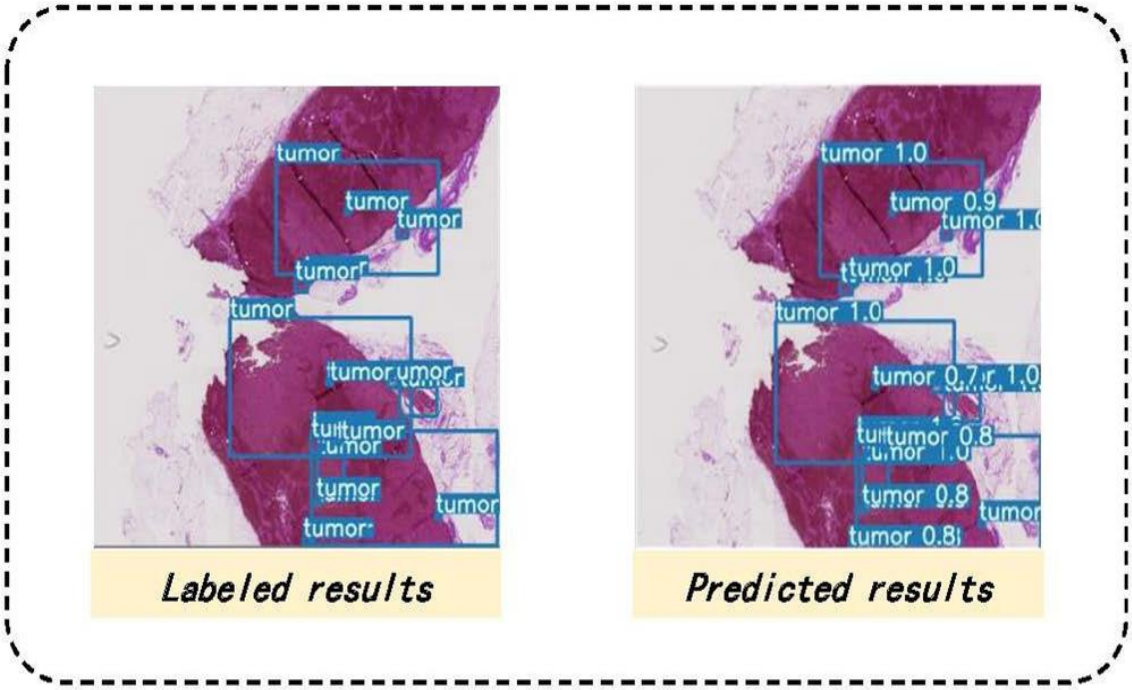
The total loss consists of two parts, the dominant loss corresponding to the output and the auxiliary loss of each sub-stage, the formula is defined as Formula 6, where denotes the weight of different loss, and denote the dominant and auxiliary loss, respectively.

The auxiliary loss branches only exist during the training stage, whereas they are abandoned when inference.

Performance of recognition module on test datasets :

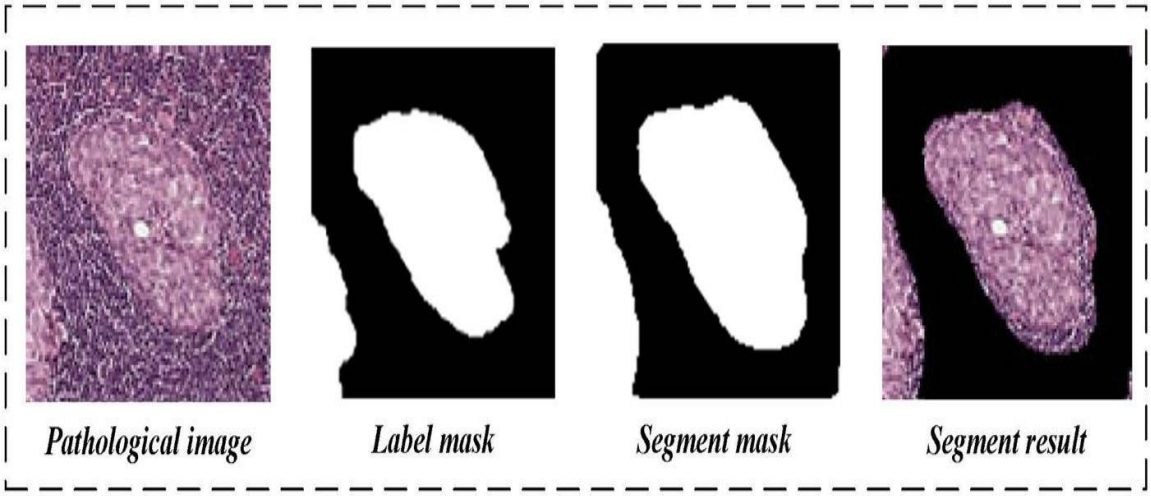
First, we used test dataset from two medical institutions to evaluate the recognition module, and the labelling and recognition results of whole slide images are shown in Figures **6** and **7**, and the evaluation results are shown in Table **2**.

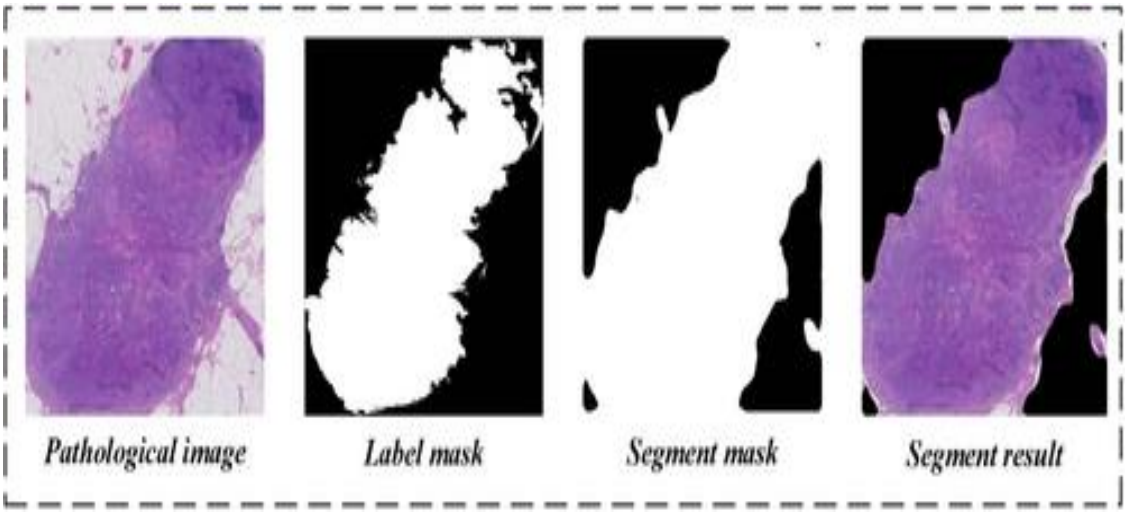




Performance of segmentation module on test datasets :

Then, we used test dataset from two medical institutions to evaluate the segmentation module, and the labelling and segmentation results of whole slide images are shown in Figures 8 and 9, and the evaluation results are shown in Table 3.





Performance display of related works :

Due to the difficult way of acquiring medical high-resolution whole slide images, most of previous works on automatic diagnosis of breast cancer pathology images used private dataset.

Moreover, different diagnostic tasks focus on different regions of interest (ROI) in the pathology image and construct own evaluation metrics based on the task.

Therefore, it is difficult to make a fair comparison of the performance of previous research works.

Although this paper uses the public dataset officially provided in the Camelyon16 competition to evaluate the model performance, the target tasks and evaluation metrics of the competition are different from this paper.

Therefore, we are unable to find suitable research works for performance comparison temporarily, and we only show the performance of related works.

By reading related papers, we found that automated diagnostic tasks for breast cancer pathology sections are divided into two main categories: recognition and segmentation.

DISCUSSION :

Most of previous works on automatic diagnosis of breast cancer pathology slide can be divided into two categories, one is to perform the task of disease classification and grading directly on low-resolution pathology slides, and the other is to crop high-resolution pathology slide into patches, then perform the task of disease classification and segmentation on the patches.

The above researches often use CNNs to extract features, then combine machine learning algorithms such as support vector machines (SVM) and ensemble learning (EM) to complete diagnostic tasks.

The results show that the diagnostic accuracy of high-resolution pathology slide is low and the diagnosis time is longer.

To solve the above problems, this study proposes a model that can quickly recognize ROI and finely segment lesion region to assist doctors in diagnosing high-resolution breast cancer pathology slide.

The proposed model achieved the recognition accuracy of 0.96 for ROI in high-resolution pathology slide, with the F1-score of 0.787; the Dice for lesion region segmentation is 0.8517.

It can be seen that the segmentation model proposed has a high recognition accuracy for ROI, and has a better segmentation effect for lesion region, which can meet the clinical diagnosis needs of pathologists.

Pathologists tend to have high false detection rate and missed detection rate when diagnosing pathology slide.

The recall of our model is 0.68, and there is also a certain false detection rate. With the development of visual transformer, compared with traditional CNNs, it can extract more local features and location information.

We will try to introduce a visual transformer into the model later, which may be useful to reduce the false detection rate of the model.

CONCLUSION :

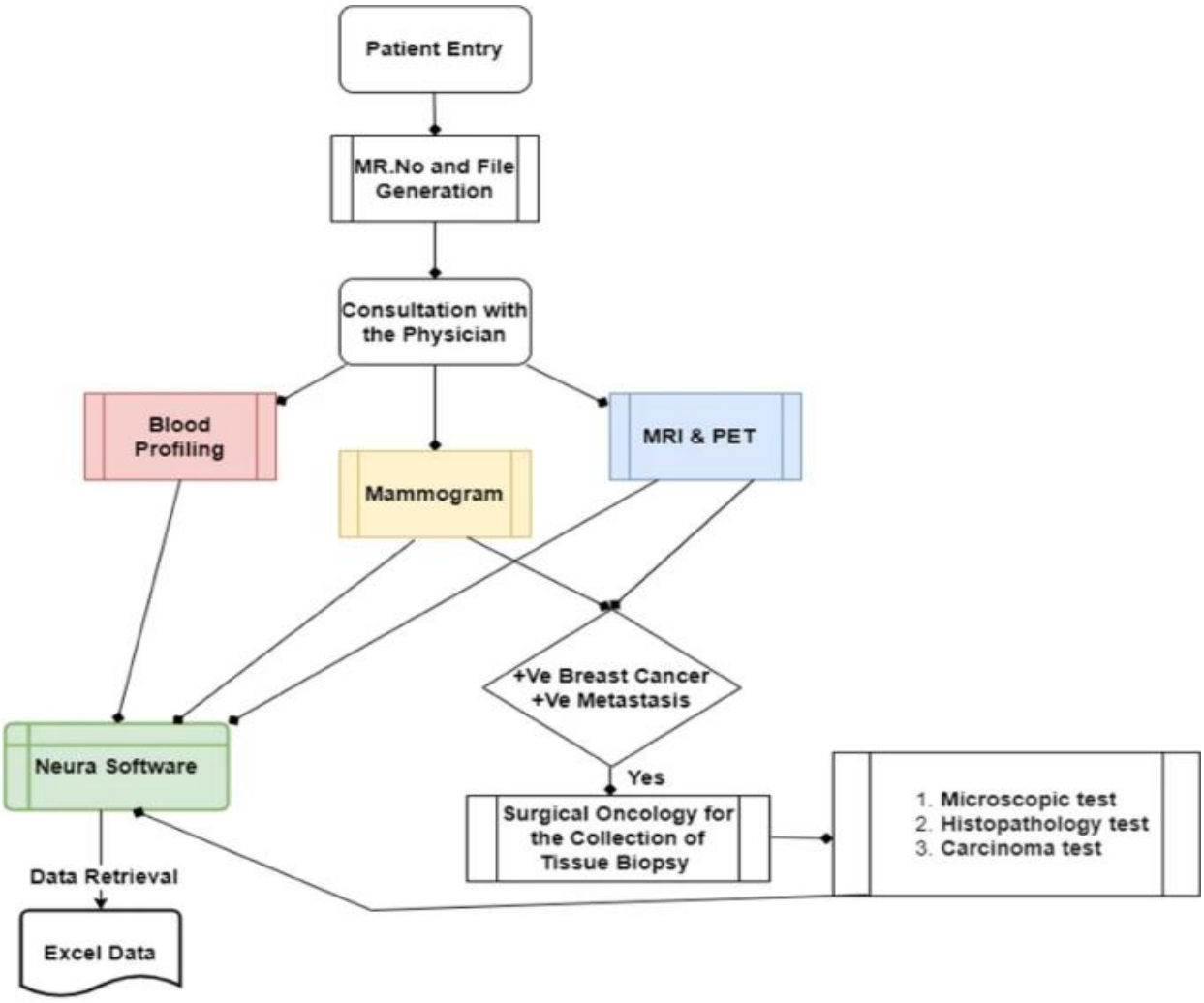
In this paper, we proposed a model to do fine segmentation of lesion region in high-resolution pathology sections, which consists of three parts involving image pre-processing module, recognition module and segmentation module.

The recognition module mainly consists of the improved single-stage object detection network YOLO-V4, that can quickly recognition the main lesion region in the pathology sections.

The segmentation module uses the global context-aware progressive aggregation network GCPANet, that can complete the fine segmentation of the lesion region.

We evaluate the proposed model on a breast cancer dataset from two medical institutions, and the experimental results show that it has a good effect on the recognition and segmentation of lesion region in high-resolution pathology sections, it can effectively improve the speed and precision of the doctor's diagnosis.

In the future, we will continue to improve the model in order to reduce the false detection rate of the main lesion region, try to apply it to the automatic diagnosis of high-resolution pathology sections of other diseases.


Patient Entry
MR.No and File Generation
Consultation with the Physician
Blood Profiling
Mammogram
MRI & PET
+Ve Breast Cancer
+Ve Metastasis
Neura Software
Yes
Surgical Oncology for the Collection of Tissue Biopsy
1. Microscopic test
2. Histopathology test
3. Carcinoma test
Data Retrieval
Excel Data