ISSN: 1813-4890

Ovarian Tumor Ultrasound Image Classification Model Based on Collaborative Attention and Improved ResNet

Yuhao Lang

School of Tianjin University of Technology and Education University, Tianjin, China;

1415750831@qq.com

Abstract

Ovarian cancer is one of the most lethal malignant tumors in the female reproductive system. Due to the lack of effective early diagnostic methods, by the time symptoms emerge, the disease has often progressed to the advanced stage. As a result, the five - year survival rate of patients with ovarian malignant tumors is low. Ultrasound is currently the most commonly used approach for the early detection of benign and malignant tumors in clinical practice. It features easy operation, non - invasiveness, and high repeatability. However, there are still two issues in the diagnostic process. Firstly, it demands high proficiency from diagnosticians, and the diagnostic results are prone to be influenced by subjectivity. Secondly, owing to the limitations of imaging technology, there may be missed diagnoses of tiny lesions, and there are certain constraints in determining the nature of tumors. In response to the above problems, an ovarian tumor ultrasound image classification model based on collaborative attention and improved ResNet is proposed. This model employs ResNet50 as the backbone network. Firstly, a collaborative attention module is added after each stage to enhance the attention to the features of low - contrast regions. Subsequently, the ordinary 3×3 convolutions in the ResNet50 network are replaced with depthwise convolutions, which not only reduces the computational burden of the model but also improves its ability to represent complex data features. After being tested with the ovarian tumor ultrasound image dataset, the experimental results show that the accuracy of the model's classification reach 91.86% respectively, effectively improving the classification results of the benign and malignant nature of ovarian tumors.

Keywords

Ovarian cancer, Collaborative attention, Image Classification.

1. Introduction

Ovarian cancer is a malignant tumor of ovarian tumors, referring to a malignant tumor that grows on the ovary. Among them, 90% - 95% are primary ovarian cancers, and the other 5% - 10% are cancers metastasized to the ovary from primary sites in other parts[1].[1] The core reason for the difficulty in early diagnosis of ovarian cancer lies in its unique anatomical characteristics and biological behavior. As a deep - seated organ in the pelvic cavity, the ovary has the characteristic of concealed growth, with a normal volume of only about 3 - 5 cm^3. Due to the lack of typical clinical manifestations, approximately 70% - 80% of patients have advanced to the late stage at the initial diagnosis. At this time, the tumor has often broken through the pelvic barrier and formed extensive implantation and metastasis in the pelvic and abdominal cavities[2]. Therefore, accurately predicting the nature of ovarian tumors and the development of lesions is of great importance, which is extremely helpful for the selection of treatment plans and prognosis assessment of patients.

ISSN: 1813-4890

Ovarian tumors include both benign and malignant tumors. Different ovarian tumors exhibit distinct clinicopathological characteristics and require different treatment methods and prognosis assessments[3]. Early - stage ovarian cancer (such as stage I and stage II) is usually mainly treated by surgical resection, with a relatively small surgical scope. After the operation, only a small amount of adjuvant chemotherapy may be required. In contrast, for advanced ovarian cancer (such as stage III and stage IV), since the tumor has spread to the abdominal cavity outside the pelvic cavity or to distant organs, the treatment plan is more complex. At the same time, multiple courses of comprehensive treatment methods such as chemotherapy, radiotherapy, or targeted therapy are also needed[4].

Ultrasound[5], due to its advantages such as convenience, economy, non-invasiveness, and non-radiation, is widely used in clinical practice. This leads to certain differences in the sensitivity and specificity of conventional ultrasound diagnosis[6]. Therefore, by means of deep learning methods, it is possible to accurately predict the nature of ovarian tumors. Benign tumors can be treated conservatively, avoiding unnecessary costs and over-treatment and preserving fertility. Malignant tumors need to be referred to the gynecological oncology department for appropriate staging and consideration of radical surgery[7]. In order to provide individualized and effective treatment plans, it is of great importance to be able to highly accurately distinguish between benign and malignant ovarian tumors.

2. Overall structure of model

2.1. DResNet structure

With the deepening of the network, during the model training process, due to the directional propagation of gradients, multi-level derivation involves successive multiplications. This may cause the gradients to approach zero, resulting in gradient vanishing, making it difficult for the network to converge or even impossible to converge. However, deepening the network also brings another problem: as the network deepens, the accuracy of both the training set and the test set decreases. Therefore, the proposal of ResNet has effectively solved the above problems.

To make the model more lightweight, in this chapter, the standard 3×3 ordinary convolution in ResNet is replaced with a 3×3 depthwise convolution. When ordinary convolution processes the input feature map, each output channel needs to perform convolution operations with all channels of the input feature map. As the network depth increases and the number of feature map channels grows, the number of parameters will increase sharply, leading to high consumption of computing resources and slower training and inference speeds. The structure of the DResNet residual block is shown in Figure1.

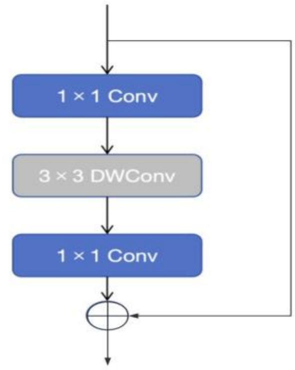

Fig. 1 DResNet residual block structure

2.2. Collaborative attention module

Channel attention reflects different information of images through each channel, which can effectively acquire important features such as echo intensity, texture and blood flow signal.

Spatial attention allows the model to focus on and around the edges, picking up subtle signs of tumor infiltration and growth. Pixel attention can analyze each pixel in the ultrasound image in detail, and reduce the problems such as low contrast and artifacts caused by the differences of ultrasound equipment and patients. The diagnostic accuracy that can be improved by a single attention is limited, so this chapter proposes a collaborative attention module CCSPA in view of the problems existing in ultrasound images and the advantages of different attention modules. It combines channel attention, spatial attention and pixel-level attention to build an effective collaborative attention mechanism, aiming to improve the model's perception ability of difficult sample features such as low contrast and shape change. The module enhances the model's focus on important regions by combining attention at different scales, and the deep convolution module enhances the modeling ability of spatial relationships. Figure 2 shows the structure.

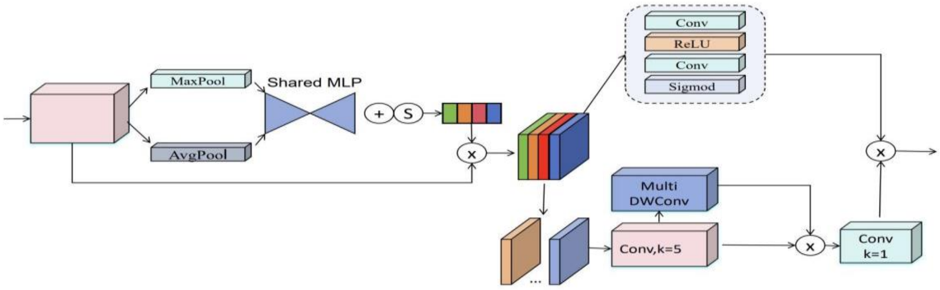

Fig. 2 Collaborative attention module structure

First, the input features are weighted by the channel attention module to obtain the enhanced channel features. In this module, the input feature maps are aggregated in channel dimension through adaptive pooling, and the results of average pooling and maximum pooling are calculated respectively.Then， the characteristics of spatial attention and pixel attention output are calculated.

$$P_{attention} = x * \sigma\left(Conv\left(ReLU\left(Conv(x)\right)\right)\right) \tag{1}$$

$$S_{attention} = Conv_{k=1}\left(\sum_k D_k(X)\right) \tag{2}$$

The final output is a weighted output of spatial attention and pixel attention, helping to enhance the
capture of globally and locally important features.

$$output = P_{attention} * S_{attention} \tag{3}$$

2.3. CAM-DResNet Model

The ultrasonic image classification model of ovarian tumor based on collaborative attention and improved ResNet proposed in this paper is based on ResNet50, and the collaborative attention module CCSPA is added after each Stage to improve the classification performance of the model. First, the input features are weighted by the channel attention module to obtain the enhanced channel features. Then, the channel enhanced features pass through the spatial attention module and the pixel attention module to further adjust the spatial and pixel-level feature representation. The final output is a weighted output of spatial attention and pixel attention, helping to enhance the capture of globally and locally important features. In order to further lighten the network model, the standard 3×3 convolution in each stage of the model is replaced by deep convolution, and the final classification results are output through global pooling and full connection layer. Figure 3 shows the overall architecture.

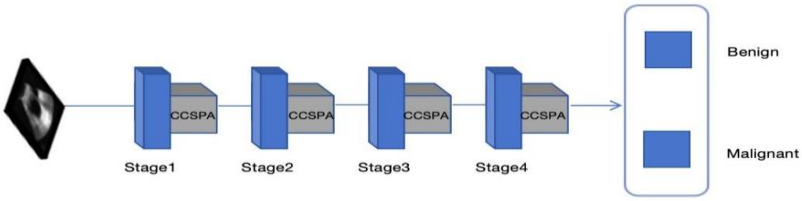

Fig. 3 CAM-DResNet model structure

3. Analysis of experimental data and results

3.1. Data sets and data preprocessing

The experimental data consisted of a total of 860 ultrasound images from clinical patients at Tianjin Medical University General Hospital, including 410 benign tumor images and 450 malignant tumor images. Initially, the ovarian cancer ultrasound image dataset was augmented by randomly flipping the ultrasound images to ensure adequate sample diversity during model training. Following data augmentation, to evaluate the model's performance, the dataset was split into a training set (80%) and a validation set (20%). This partitioning process ensured both the randomness of the samples and a balanced distribution across different classes, thereby preventing potential bias during model training and validation. During the training process, the Adaptive Moment Estimation (Amdam) optimizer is used to accelerate the training and enhance the generalization ability of the model for the clinical complex scenes. The initial learning rate is set to 0.001.

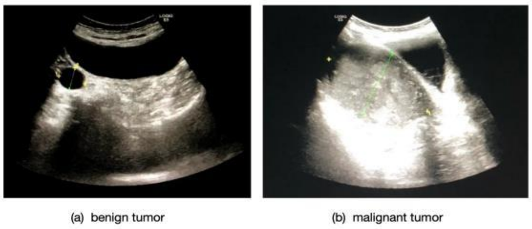

Fig. 4 Benign and malignant tumors

3.2. Analysis of experimental results

In order to verify the performance of the CAM-DResNet model proposed in this chapter, we compared it with the deep learning classical classification networks AlexNet, VGG16, GoogleNet, MobileNet, ResNet and DenseNet[8]. The experimental results are shown in Table 1.

Table 1 Experimental results of each model

Models	Accurary	Specificity	Precision	Recall	F1
AlexNet	0.8167	0.7679	0.8034	0.7953	0.8284
VGG16	0.8247	0.8137	0.7910	0.8159	0.7823
GoogleNet	0.8628	0.8392	0.8367	0.8597	0.8481
MobileNet	0.8541	0.8461	0.8671	0.8463	0.8407
ResNet34	0.8794	0.8573	0.8861	0.8618	0.8859
ResNet50	0.8827	0.8687	0.8972	0.8856	0.8914
DenseNet	0.8781	0.8569	0.8957	0.8543	0.8769
CAM-DResNet	0.9186	0.8873	0.9184	0.9046	0.9137

Experimental results show that the classification accuracy of ResNet50 is higher than that of ResNet34, so this chapter uses ResNet50 as the backbone network to achieve better classification effect. Compared with the backbone model ResNet50, the Acc of CAM-DResNet increased by 3.59%. As for the ability of the model to correctly identify negative samples, the Spe of CAM-DResNet was 88.73%, higher than that of GoogleNet (83.92%) and MobileNet

(84.61%). Considering the F1 index of Precision and Recall comprehensively, the F1 value of CAM-DResNet was 91.37%, which was much higher than 78.23% of VGG16 and 84.07% of MobileNet.

In order to verify the superiority of DResNet and collaborative attention module CAM proposed in this chapter, experiments were carried out to verify ResNet, DResNet, CAM-ResNet and CAM-DResNet[9]. The accuracy of model classification and the number of model parameters were shown in Table 2.

Table 2 The result of different module combinations

Models	Accurary	Parameter/M
ResNet	0.8798	21.57
DResNet	0.8837	11.48
CAM-ResNet	0.9079	24.16
CAM-DResNet	0.9186	14.39

As can be seen from Table 2, the classification accuracy of CAM-DResNet proposed in this chapter is the highest, reaching 91.86%, and the number of parameters is significantly reduced compared with traditional networks. When the collaborative attention module is directly embedded into ResNet, its classification accuracy is also improved by 2.81%, which proves that the collaborative attention module proposed in this paper can well capture global and local information and improve the model's attention to important features.

4. Conclusion

The ultrasonic image classification model of ovarian tumor based on collaborative attention and modified ResNet proposed in this chapter improves the accuracy of the model classification by addressing factors such as low contrast of ultrasound images and uncertain location and size of lesions. Deep separable convolution can reduce the number of parameters in the model, and the collaborative attention module can highlight focal area information and capture detailed features. It can be seen from the experimental data that compared with traditional ResNet, its classification accuracy is increased by 3.88%, the number of parameters is reduced by 7.18M, and it has good classification performance.

References

[1] Di Wen, Hu Yuan. Big data study on ovarian cancer [J]. Chinese Journal of Practical Gynecology & Obstetrics,2018,34(01):18-22.

[2] Wei Mingxiang; Cypress foundation; Guo Lili; Research progress of imaging omics in ovarian tumors [J]. Magnetic Resonance Imaging,2020(05):76-79.

[3] Chen Huizhu,WANG Mengtian, Ning Gang. Advances in imaging diagnosis of ovarian malignancy [J]. Chinese Journal of Maternal and Child Clinical Medicine (Electronic edition), 2015, 11 (01): 90-94.

[4] Tian Zhongyan; [J]. Study on ultrasonic examination to improve the diagnosis rate of ovarian tumor [J]. Chinese Urban and Rural Enterprise Health,2019(06):97-98.

[5] Mukaida S H ,Matsuzaki S ,Ueda Y , et al. Systematic Review of the Survival Outcomes of Neoadjuvant Chemotherapy in Women with Malignant Ovarian Germ Cell Tumors [J]. Cancers, 2023, 15 (18):78-82

[6] Kemppainen J ,Hynninen J ,Virtanen J , et al. PET/CT for Evaluation of Ovarian Cancer [J]. Seminars in Nuclear Medicine, 2019, 49 (6): 484-492.

[7] Yang Yan, Zhang Guomei, Li Hongjuan, et al. The value of ultrasonography, CT and MRI combined with serum CA125 in determining the nature of ovarian tumor [J]. Journal of Practical Cancer, 2021, 36 (07): 1144-1147

[8] Srivastava S, Kumar P, Chaudhry V, Singh A. Detection of Ovarian Cyst in Ultrasound Images Using Fine-Tuned VGG-16 Deep Learning Network. SN Comput Sci (2020) 1:81.

[9] Jia W ,Gahb S J . Research on Local Feature Intelligent Extraction Algorithm of Blurred Image Under Complex Illumination Conditions [J]. IEEE ACCESS, 2021, 9 84948-84955.