Key Clinical Parameters Detection and Ovarian Tumor Benign/Malignant Classification in Multi-Modal Ultrasound Images via a Multi-Task Model

Chunjun Qian, Lulu He, Xianglian Meng, Mengxia Yu, Xiuhua Wu, Yanyun Shi (shiyanyun2009@aliyun.com)

Abstract. Ovarian tumors (OT) are one of the most common tumors in the female reproductive system, and are often classified into two categories: benign and malignant. B-mode ultrasound (BMUS) and contrast-enhanced ultrasound (CEUS) are commonly used as conventional imaging techniques for ovarian tumor screening; however, image interpretation is time-consuming. In this study, we construct a multi-task multi-modal ultrasound image dataset. A multi-task model based on a pre-trained deep learning architecture is trained and applied to predict and select key clinical parameters, as well as to classify benign and ovarian tumors. Finally, blood flow signal, morphology, spetation, and solid component are selected as key clinical parameters related to the classification of benign and malignant ovarian tumors. The evaluation metrics for classifying benign and malignant ovarian tumors are as follows: accuracy 0.731, precision 0.692, recall 0.643, and F1-score 0.653. The method based on deep learning and medical imaging has the potential to assist in accelerating the initial screening of benign and malignant ovarian tumors and generating image-to-text inspection reports.

Keywords: Ovarian Tumor Classification, Multi-Task Model, Multi-Modal Ultrasound Image.

1 Introduction

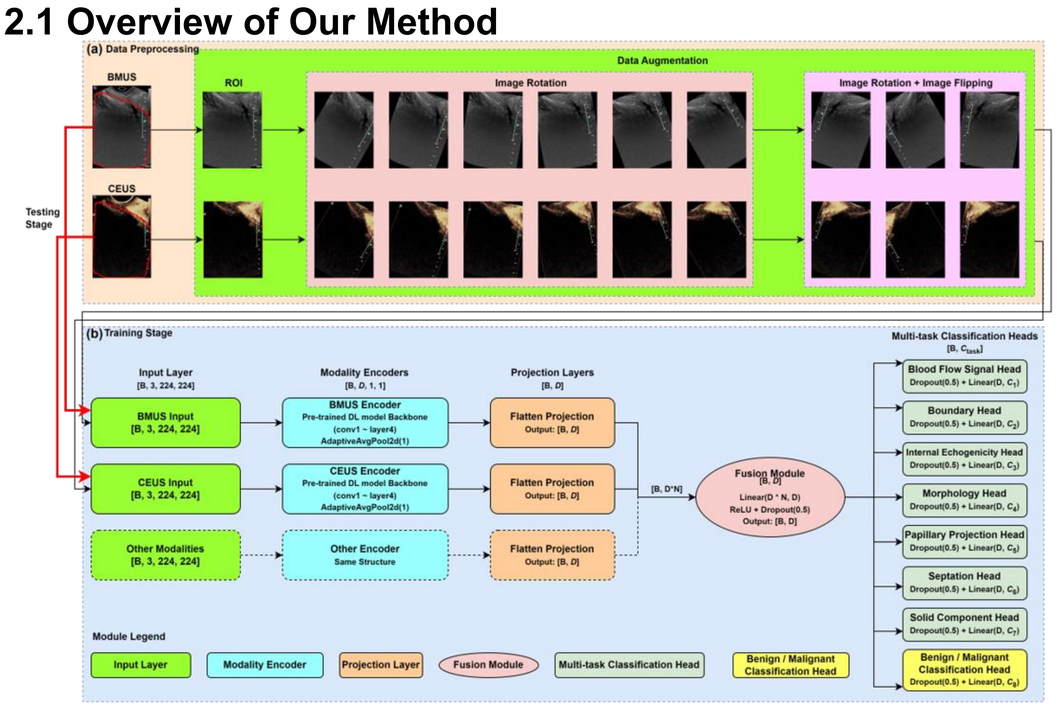
Ovarian tumors (OT) are one of the most common tumors in the female reproductive system and can occur at any age. They are classified into two categories: benign and malignant [1]. Ultrasound imaging is the primary imaging method for ovarian tumors [3]. The clinicians can use the ultrasound imaging results from radiologists to determine whether an ovarian tumor is benign or malignant preliminarily. However, ultrasound imaging reports, which are prepared by radiologists and involve recording and interpreting ultrasound images, require significant time and effort [4].

To distinguish between benign and malignant ovarian tumors, researchers have developed and evaluated the Ovarian-Adnexal Reporting and Data System (O-RADS), a system based on BMUS analysis [9,10]. Recently, deep learning (DL) approaches [13-17] have been introduced for the automated classification of benign and malignant ovarian tumors. In practice, BMUS or CEUS examinations for characterizing ovarian tumors aim to provide, for a given image or image set, both benign/malignant prediction and the selection of key clinical parameters (e.g., solid components, morphological features, and blood flow signals).

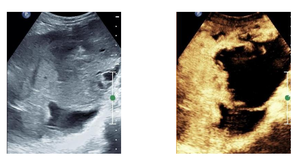
To address this gap, we propose constructing a multi-label, multi-modal ovarian tumor dataset and establishing a corresponding prediction model based on pre-trained deep learning architectures. This model will simultaneously output key clinical parameter predictions and classify ovarian tumors as benign or malignant. This integrated approach reduces model complexity, the number of required models, and the magnitude of model parameters, thus conserving significant computational resources and improving processing speed.

2 Method

2.1 Overview of Our Method



2.2 Dataset Construction



	Key Clinical Parameters																Classification	
	Blood flow signal		Boundary		Internal echogenicity				Morphology		Papillary projection		Septation		Solid component		-	
Description	Non-exist	Exist	Clear	Unclear	Anechoic	Hypoechoic	Isoechoic	Hyperechoic	Regular	Irregular	Non-exist	Exist	Non-exist	Exist	Non-exist	Exist	Benign	Malignant
Label	0	1	0	1	0	1	3	4	0	1	0	1	0	1	0	1	0	1
Sample		√		√				√		√	√		√			√		√

2.3 Multi-Task Multi-Modal Classification Method

The multi-task multi-modal classification problem is formulated as $Y = f(x_{BMUS}, x_{CEUS})$, where $Y = \{y_1, y_2, \cdots, y_{C_{task}}\}$ represents the labels of multi-task C_{task} for a given pair of BMUS x_{BMUS} and CEUS x_{CEUS} images, and f means the mapping operations, such as convolution, image transformation, fully connected layer, softmax, or sigmoid output.

For each task, the loss function can be formulated as

$$L_{C_{task}} = -\sum_{i=1}^{C_{task}} y_{i_{task}} \log \hat{y}_{i_{task}}$$

The total loss function can be defined as

$$L = \sum_{task=1}^{8} L_{C_{task}} = -\sum_{task=1}^{8} \sum_{i=1}^{C_{task}} y_{i_{task}} \log \hat{y}_{i_{task}}$$

3 Experimental Results

3.1 Datasets

Parameter \ Category	Benign				Malignant			
	'0'	'1'	'3'	'4'	'0'	'1'	'3'	'4'
Blood flow signal	110	64	-	-	6	77	-	-
Boundary	170	4	-	-	63	20	-	-
Internal echogenicity	109	29	5	31	13	18	8	44
Morphology	152	22	-	-	36	47	-	-
Papillary projection	155	19	-	-	70	13	-	-
Septation	103	71	-	-	71	12	-	-
Solid component	141	33	-	-	42	41	-	-

3.2 Evaluation Metrics

Accuracy, precision, recall, and F1-score

3.3 Results

Table 2. Evaluation metrics of distinguishing benign/malignant ovarian tumor with different pre-trained deep learning models (bold font means the best result).

Method \ Evaluation metric	Accuracy	Precision	Recall	F1-score
EfficientNet	0.667	0.590	0.565	0.563
MobileNetV2	0.641	0.549	0.535	0.529
MobileNetV3_small	0.718	0.671	**0.634**	**0.642**
MobileNetV3_large	**0.731**	**0.696**	0.633	0.641
VGG16	0.654	0.507	0.502	0.457
VGG19	0.705	0.660	0.582	0.576
ResNet18	0.705	0.688	0.561	0.537
ResNet34	0.628	0.508	0.505	0.486
ResNet50	0.654	0.574	0.555	0.553
DenseNet121	0.692	0.631	0.573	0.567
DenseNet169	0.692	0.633	0.562	0.549
DenseNet201	0.692	0.631	0.583	0.583
ViT	0.705	0.660	0.582	0.576

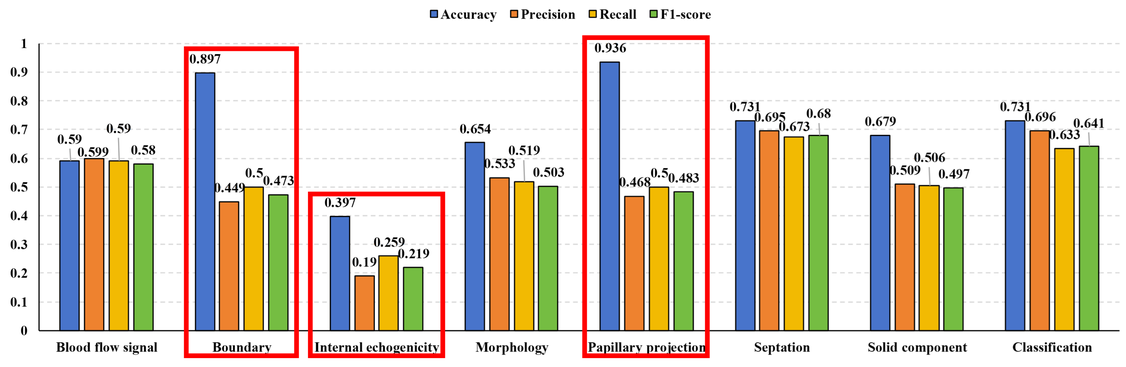


Fig. 2. Evaluation metrics for each clinical parameter with MobileNetV3_large. The clinical parameters outlined in the red rectangle are not considered in this work.

Table 3. Evaluation metrics for selected clinical parameter predictions and classification of benign and malignant ovarian tumors.

Multi-task \ Evaluation metric	Accuracy	Precision	Recall	F1-score
Blood flow signal	0.628	0.658	0.628	0.610
Morphology	0.654	0.514	0.507	0.481
Spetation	0.628	0.565	0.558	0.558
Solid component	0.705	0.516	0.507	0.485
Classification	0.731	0.692	0.643	0.653

Reference

[1] D. Timmerman, et al. Logistic Regression Model to Distinguish Between the Benign and Malignant Adnexal Mass Before Surgery: A Multicenter Study by the International Ovarian Tumor Analysis Group. Journal of Clinical Oncology (2005) 8794-8801.
[3] D. Fischerova, et al. Ultrasound in preoperative assessment of pelvic and abdominal spread in patients with ovarian cancer: a prospective study. Ultrasound in obstetrics & gynecology (2017) 263-274.
[4] W. Dai, et al. Development and validation of a deep learning pipeline to diagnose ovarian masses using ultrasound screening: a retrospective multicenter study. EClinicalMedicine (2024).
[9] L. Shen, et al. The Ovarian-Adnexal Reporting and Data System (O-RADS) US Score Effect on Surgical Resection Rate. Radiology (2024) e240044.
[10] Y. Yi, et al. Diagnostic performance of contrast-enhanced ultrasound (CEUS) combined with Ovarian-Adnexal Reporting and Data System (O-RADS) ultrasound risk stratification for adnexal masses: a systematic review and meta-analysis. Clinical Radiology (2024) e1167-e1175.
[13] T.-L. Pham, V.-H. Le. Ovarian tumors detection and classification from ultrasound images based on YOLOv8. Journal of Advances in Information Technology (2024).
[14] W. Xie, et al. Developing a deep learning model for predicting ovarian cancer in Ovarian-Adnexal Reporting and Data System Ultrasound (O-RADS US) Category 4 lesions: A multicenter study. J Cancer Res Clin Oncol (2024).
[15] Y. Du, et al. Development and validation of an ultrasound-based deep learning radiomics nomogram for predicting the malignant risk of ovarian tumours. BioMedical Engineering OnLine (2024).
[16] Y. Du, et al. Ultrasound-based deep learning radiomics model for differentiating benign, borderline, and malignant ovarian tumours: a multi-class classification exploratory study. BMC Medical Imaging (2024).
[17] S. Zeng, H. Jia, H. Zhang, et al. Multimodal ultrasound-based radiomics and deep learning for differential diagnosis of O-RADS 4-5 adnexal masses. Cancer Imaging (2025).

Acknowledgments

This study was funded by the National Natural Science of China (Grant NO. 62201095, 62370174), Science and Technology Projects of Changzhou Municipal Health Commission (Grant NO. ZD202341), and Changzhou Science and Technology Plan (International Science and Technology Cooperation) (Grant NO. CZ20230028).