Breast Cancer Classification through Meta-Learning on Multimodal MRIs

Raymond Hon-Fu Chan
Dept. of Operations and Risk Management
Lingnan University of Hong Kong
Hong Kong, China
rhfchan@ln.edu.hk

Yao Lu
School of Computer Science and Engineering
Sun Yat-sen University
Guangzhou, China
luyao23@mail.sysu.edu.cn

Chenghao Qiu
Dept. of Mathematics
City University of Hong Kong
Hong Kong, China
chenghqiu2-c@my.cityu.edu.hk

Jun Xie
Dept. of Mathematics
City University of Hong Kong
Hong Kong, China
junxie9@cityu.edu.hk

Siu Pang Yung
Dept. of Mathematics
City University of Hong Kong
Hong Kong, China
spyung@hku.hk

Kehui Zhang
Dept. of Mathematics
City University of Hong Kong
Hong Kong, China
kehuzhang2@cityu.edu.hk

Xiaosheng Zhuang
Dept. of Mathematics
City University of Hong Kong
Hong Kong, China
xzhuang7@cityu.edu.hk

***Abstract*—This paper proposes a multimodal neural network AI model for gauging the metastatic load of axillary lymph nodes in the breast. The model utilizes three modalities of images, namely dynamic contrast enhancement (DCE), T2-weighted (T2W), and diffusion-weighted imaging (DWI), from breast magnetic resonance imaging (MRI) and axillary lymph node MRI. Features are extracted by a feature extractor (composed of conv1 and layer1 of ResNet and Wavelet transform convolution model) that is pretrained on a large breast cancer MRI dataset based on the Model-Agnostic Meta-Learning (MAML) algorithm. The features of the same modality from breast MRI and axillary lymph node MRI are concatenated and then input into the multimodal MulT model for classifications. The experimental results show that the addition of meta-learning and the involvement of multimodal MRI (rather than just uni-model MRI) significantly improve the classification, with the area under the ROC curve (AUC) reaching 0.84. The model performs well in judging the metastatic load of axillary lymph nodes in the breast and is expected to contribute to clinical diagnosis and treatments (both invasive and non-invasive).**

***Index Terms*—multimodal MRI, breast cancer classification, meta-learning, deep learning, medical imaging diagnostic analytics**

I. Introduction

Breast cancer has been the most frequently diagnosed cancer among women globally, with axillary lymph node metastasis (ALNM) being a key factor that influences prognosis and treatment choices [1]. According to current clinical recommendations, sentinel lymph node biopsy (SLNB) is considered the standard for assessing nodal involvement [2]. However, research indicates that between $20-35\%$ of SLNBs produce false negatives, even when preoperative imaging is conducted. These inaccuracies can lead to unnecessary surgical complications, such as lymphedema and sensory neuropathy [3]. This issue highlights the urgent need for improved non-invasive methods to accurately detect hidden metastases, particularly micrometastases $\leq 2mm$, which often go undetected by standard procedures.

Multimodal magnetic resonance imaging (MRI) is widely used for the screening, diagnosis, and treatment of breast cancer [4]. Despite the availability of various modalities, including Dynamic Contrast-Enhanced Imaging (DCE), T2-weighted Imaging (T2W), and Diffusion-Weighted Imaging (DWI), there are still three major challenges in detecting axillary lymph node metastasis (ALNM): (1) Cross-regional biological correlation modeling: The interaction between primary tumors and lymph nodes presents spatially diverse characteristics, such as tumor angiogenesis from DCE compared to cellularity observed in DWI, which complicates conventional registration techniques. A significant spatial mismatch ($> 5mm$ in 38% cases) and differing contrast mechanisms make meaningful feature pairing difficult [5]. (2) The Prostate Imaging-Reporting and Data System (PI-RADS) stipulates that DCE imaging must be utilized to further classify peripheral zone (PZ) cases with a DWI score of 3 (DWI3). A positive DCE score can elevate the overall assessment to a 4, indicating clinically significant prostate cancer (csPCa). Moreover, DCE enhances the accuracy of predicting csPCa in DWI3 PZ instances [5]. Current fusion techniques often overlook specific enhancement patterns that are crucial for identifying micro-metastases. These challenges diminish clinical confidence in AI-assisted nodal staging and increase patients' dependence on invasive biopsies.

Meta-learning approaches generally involve a meta-training stage where a series of few-shot tasks generated from base classes are used during the training period. The strategy enables well-crafted models to rapidly adjust to new, unseen tasks during testing [6]. Essentially, the meta-training phase is composed of two nested loops. Initially, the model parameters are updated using the training samples within each task. Subsequently, the model undergoes optimization by meta-fine-tuning on the test samples of each task. Through this process, the meta-learning framework acquires meta-knowledge spanning multiple tasks, which facilitates swift adaptation to novel tasks.

Although previous studies highlight individual strengths [7], they primarily focus on analyzing various MRI modalities separately, overlooking the biological interactions across different regions. On the other side, traditional implementations of ResNet [8] and wavelet transform convolution (WTConv) [9] face challenges due to scanner-specific domain variations, an issue that our proposed meta-learning enhanced framework aims to tackle directly.

Given these challenges, this paper seeks to develop an efficient and precise classification model for assessing the metastatic load in breast lymph nodes. We enhance both the accuracy and dependability of classifying the metastatic burden by using advanced feature extraction methods and meta-learning techniques. This would offer strong support for related clinical diagnosis and treatment of breast cancer. The specific areas of research are outlined as follows:

- **Research on the Feature Extractor:** In the initial phase of our model, we will thoroughly utilize the feature extractor based on ResNet18 [8] and WTConv [9]. Through using ResNet18's strengths in image feature extraction alongside with the distinctive convolutional capabilities of WTConv, we can more effectively identify critical features in breast lymph node images. We aim at ensuring that the feature extractor accurately captures the information pertinent to the metastatic load of lymph nodes by carefully designing the network architecture and optimal parameter configurations. This lays a strong groundwork for the classification tasks that follow.
- **Research on Meta-learning Pre-training Method:** Implementing the meta-learning approach to pre-train the feature extractor is a central focus of this study [6]. Meta-learning allows the model to acquire overarching knowledge and strategies across a range of tasks, enabling it to adapt effectively to new challenges. In our experiment, we conduct meta-learning and pre-train it on diverse medical image datasets. The feature extractor will construct generalized representations of image features by refining meta-learning tasks, such as contrastive learning and self-supervised learning. This enhances the ability of the model to generalize, and consequently, it will be better equipped to process medical images from various sources with differing characteristics, yielding more universal features for assessing the metastatic burden in breast lymph nodes.
- **Multimodal Feature Extraction and Fusion:** Our research centers on two primary anatomical areas: the breast and axillary lymph nodes. We analyze MRI images across three modalities—DCE, T2W, and DWI—from both regions. We will employ pre-trained feature extractors to carry out feature extraction for images from these distinct anatomical areas and modalities independently. Despite sharing the same architecture and pre-trained weights, each feature extractor will analyze its respective images separately, allowing for an in-depth examination of the unique information present in each anatomical region and modality. Following this, we will develop innovative and effective fusion strategies to combine features across these areas and modalities. During the fusion process, we will integrate the features obtained from the breast and axillary lymph nodes across different modalities, fully uti-

lizing the complementary strengths of multimodal information to enhance the accuracy of classifying the metastatic burden of breast lymph nodes. Our approach of feature fusion will provide a more comprehensive depiction of lymph node characteristics from multiple perspectives, effectively overcoming the limitations associated with relying on a single anatomical region or modality. Our fusion framework will utilize MulT [20], which offers a distinct advantage in feature fusion through its cross-modal attention mechanism. This model can capture long-range dependencies across modalities, thereby addressing gaps between different modal data while capturing long-range dependencies across modalities. This allows for the effective use of multimodal information without the need for explicit alignment, and enhances the fusion of features extracted from various MRIs.

In conclusion, the main contributions of this paper are outlined below:

- Create and deploy a pre-trained feature extractor utilizing the MAML algorithm. Integrate conv1 and layer1 of ResNet with the WT-conv model to efficiently extract features from breast MRI and axillary lymph node MRI.
- Utilize a multimodal (MulT) model to integrate features from various modalities derived from breast MRI and axillary lymph node MRI. This enable us to thoroughly examine the complementary information present in multimodal data, facilitating a precise classification of the metastatic load in breast axillary lymph nodes.
- Perform comparative experiments to validate the effectiveness of meta-learning and the fusion of multimodal data in enhancing the model's performance.
- Train and evaluate the model using actual datasets of breast MRI and axillary lymph node MRI, assessing its classification performance based on the AUC metric.

II. Related Work

Meta-learning has seen good development in image classification in recent years. In 2017, Finn *et al.* [6] proposed Model-Agnostic Meta-Learning (MAML), a meta-learning method based on optimization. The main idea of MAML is to learn the initialization of a neural network so that it follows the fast gradient direction, thus effectively classifying new classes. Prototypical networks for few-shot learning [10] is a meta-learning method based on metric learning, which learns the embedding of the dataset into a low-dimensional space and classifies the dataset through cosine similarity. These methods have shown good performance on some common few-shot datasets, such as $mini$ImageNet [11] or Omniglot [12]. Learning to learn by gradient descent by gradient descent [13] is a meta-learning method for optimizing the optimizer. Different from previous methods, learning to learn by gradient descent by gradient descent [14] considers using meta-learning to optimize the parameters of the optimizer, so as to improve the learning performance of the neural network. Bayesian meta-learnings were also proposed so as to learn the probability distribution of image classification, and use the evidence lower bound (ELBO) as the optimization function or approximate the probability distribution using MAP [14] for learning.

In comparison, there are relatively few studies on few-shot learning of medical images. Mahajan *et al.* [16] implemented the recognition of a small number of skin diseases under the setting of long-tailed class distribution based on the meta-learning framework and attempted rapid model adaptation. Hu *et al.* [17] designed a novel data augmentation method, which was not in the input space but in the logit space, effectively alleviating the overfitting problem of classification tasks with limited medical images. Also, [17] formulated the few-shot learning problem of retinal diseases as a student-teacher learning task, and simultaneously utilized the discriminative feature space and Knowledge Distillation (KD) technology. In this paper, we shall propose a few-shot learning classification method for medical images based on MAML, which integrates the advantages of both fine-tuning and meta-learning.

III. Methodology

A. Problem Formulation and Method Overview

Given paired inputs $\{X_1, X_2, X_3, X_4, X_5, X_6\}$, where X_1, X_2, X_3 represent breast MRI scans across three modalities (DCE, DWI, T2W), and X_4, X_5, X_6 denote corresponding axillary lymph node MRI scans for the same modalities, our goal is to predict a binary classification y (metastatic vs. non-metastatic). To address this, we propose a novel framework that first employs modality-specific feature extractors to encode each input $X_i (i = 1, \ldots, 6)$. Features from anatomically aligned modality pairs (e.g., breast DCE X_1 and lymph node DCE X_4) are concatenated to preserve cross-regional biological interactions. These fused modality features are then integrated through a hierarchical multimodal transformer, which learns inter-modal dependencies between DCE (vascular kinetics), DWI (cellularity), and T2W (morphology). Finally, the model predicts y based on the joint representation, optimizing for both metastatic sensitivity and specificity. The pipeline—spanning feature extraction, cross-region fusion, and multimodal integration—is illustrated in Fig. 1, with color-coded pathways ensuring clarity in modality-specific and cross-modal information flow.

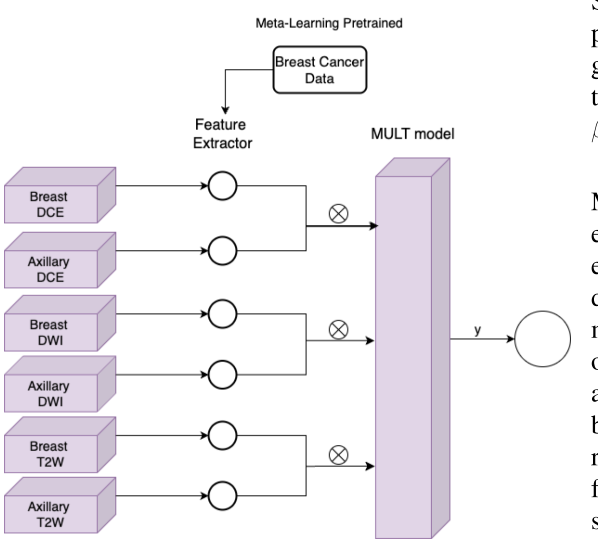

Fig. 1. The multimodal MRI represented by the six blocks, as shown on the left side, are encoded separately in the meta-learning pre-trained feature extractors. The features obtained from the images of the same MRI modality are concatenated pairwise and fed into the multimodal Transformer model [20], and the classification results are obtained.

B. Pre-training Process Based on MAML

The researchers developed a pre-training framework integrating ResNet18 and WTConv, optimized through the MAML algorithm, to empower the feature extractor with generalized image representation capabilities and rapid adaptability across diverse medical imaging tasks. We initialized the parameters of the network randomly and optimized them using meta tasks. These tasks are obtained from BC-data and structured under meta-learning principles, with each task containing a dedicated support set and query set to ensure clinically relevant feature extraction.

This feature extractor addresses the transferability limitations of conventional medical imaging models by fostering both broad feature abstraction and task-sensitive refinement. The innovation of the framework lies in its two-stage optimization architecture: meta pre-training learns foundational feature representations, and training on specific tasks equips the model with better feature extraction ability for specific tasks, effectively bridging general and domain-specific medical image analysis.

For each sampled task, gradient descent is performed on the support set using the current initialized parameters, and the parameters are updated according to the formula $\theta_i' = \theta - \alpha \nabla_\theta L_{T_i}(\theta)$ to adapt to the task, where α is the learning rate. Subsequently, the query set is used to evaluate the performance of the model and calculate the meta-gradient $\nabla_\theta L_{T_i}(\theta_i')$. Then, the initialized parameters are updated according to the meta-learning rate β and the formula $\theta = \theta - \beta \nabla_\theta \sum_{i=1}^{n} L_{T_i}(\theta_i')$.

The efficiency of the framework is due to MAML's optimization strategy. MAML strengthens the task-agnostic adaptability of the model, enabling efficient parameter fine-tuning on minimal data to achieve robust performance across unseen medical image tasks. There exist two advantages of MAML: the feature extractor not only generalizes foundational patterns from prior knowledge but also rapidly reconfigures its hierarchical representations to capture task-specific discriminative features, which can obtain a good performance by small datasets.

Critical to this process is the algorithmic emphasis on meta-optimization. By exposing the model to probabilistically sampled task distributions during pre-training, the architecture develops intrinsic sensitivity to subtle feature variations across modalities. The support-query paradigm further reinforces this capability, guiding the model to distill transferable representations while avoiding overfitting to ephemeral task characteristics.

Algorithm 1 Model-Agnostic Meta-Learning

$p(\mathcal{T})$: distribution over tasks;
α, β: step size hyperparameters
• randomly initialize θ;
• while not done:
sample batch of tasks $\mathcal{T}_i \sim p(\mathcal{T})$,
• for all $\mathcal{T}_i$ do:
evaluate $\nabla_\theta \mathcal{L}_{\mathcal{T}_i}(f_\theta)$ with respect to K examples;
compute adapted parameters with gradient descent: $\theta_i' = \theta - \alpha \nabla_\theta \mathcal{L}_{\mathcal{T}_i}(f_\theta)$;
End for
• update $\theta \leftarrow \theta - \beta \nabla_\theta \sum_{\mathcal{T}_i \sim p(\mathcal{T})} \mathcal{L}_{\mathcal{T}_i}(f_{\theta_i'})$.

C. Feature Extractor

For breast and axillary lymph node MRI images in DCE, T2W, and DWI modalities, we deploy a meta-learning pre-trained ResNet-based feature extractor. Comprising layer1, the initial convolutional part, and WTConv [18], [19], the identically-structured extractors for each modality operate independently. They adapt to anatomical and modal differences, unearthing unique features despite sharing pre-trained weights.

For breast DCE-MRI images, we first resize them to 224×224 and normalize. The images then pass through the initial 7×7 convolutional layer with a stride of 2. This downsamples the images to 112×112 and increases channels to 64. A 3×3 max-pooling layer further reduces resolution to 56×56, keeping channels at 64. Next, the images enter two BasicBlocks in layer1. Each BasicBlock has two 3×3 convolutional layers, followed by BN and ReLU. The first layer captures local features, and the second extracts complex ones. Without downsampling in layer1 BasicBlocks, both convolutional layers have a stride of 1, maintaining 56×56 size and 64 channels for low-level feature extraction.

Subsequently, images enter the WTConv module. Wavelet transform decomposes images into different-frequency components. Small convolutional kernels then capture local features at each frequency: low-frequency for shapes and structures and high-frequency for details. After convolution, the inverse wavelet transform recombines components into feature maps. These rich-feature outputs provide important information for subsequent network layers.

D. Cross-anatomical Region and Multimodal Integration Module

This study introduces a hierarchical multimodal fusion framework that synergizes DCE, T2W, and DWI MRI data through three interconnected computational stages to optimize breast lymph disease diagnosis. Initially, anatomically specific features from breast and axillary lymph node regions are extracted in parallel using pre-trained encoders for each modality, with spatial concatenation creating cross-region unified representations that preserve pathological correlations. These intramodal features then undergo cross-modal integration via a Multimodal Transformer (MulT) architecture, where bidirectional attention mechanisms autonomously model diagnostic-relevant interactions—such as hemodynamic linkages between DCE and T2W sequences—without requiring explicit data alignment. The final stage enhances clinical interpretability through diagnostic-adaptive feature refinement, resolving contradictory multimodal signals (e.g., discordant DWI/T2W intensities) via gated weighting while dynamically correlating lesion enhancement patterns with nodal metabolic characteristics.

IV. Experiment

A. Experimental Setup

1) Dataset: Breast cancer MRI and axillary lymph MRI Dataset contains 1572 from 2 classes (567 low metastatic burden of axillary lymph nodes in the breast and 1005 high metastatic burden of axillary lymph nodes in the breast). The number of DCE-MRI is 524. The number of DWI-MRI is 524. The number of T2W-MRI is 524.

2) Baseline: We consider cases of uni-modal (DCE-only, DWI-only and T2W-only), non-meta-learning multi-modality, and meta-learning multi-modality for comparison.

a) *Meta-learning Ablation Experiment*: To verify the role of meta-learning in the model, a meta-learning ablation experiment was conducted. Two groups of experiments were set up. In one group (the experimental group), the model was trained using the feature extractor pre-trained with the MAML algorithm. In the other group (the control group), the feature extractor was directly trained on the breast MRI and axillary lymph node MRI datasets. The performance of the models in the two groups was compared to reveal the impact of meta-learning on the model performance. The feature extractor we use is a combination of ResNet18.conv1, ResNet18.layer1 and WTConv. We train all of them in the meta-learning pretrained part, and we froze ResNet.conv1 during the training.

b) *Uni-modal Experiment*: To verify the performance improvement of the model brought by multi-modal MRI, a uni-modal experiment was carried out. The DCE, T2W, and DWI modal images of breast MRI and axillary lymph node MRI were used separately to train the model. Three uni-modal models were obtained and their performance were compared to that of the multi-modal model to reveal the performance improvement effect of multi-modal fusion on the model.

c) *Multi-modal Experiment:* To verify the impact of replacing certain modules on the performance of the multi-modal model, a series of multi-modal experiments were conducted. In the original multi-modal model that integrates modalities such as DCE, T2W, and DWI, modules like ResNet.layer1 and WTConv were replaced separately. For example, ResNet.layer1 was replaced with a newly designed convolutional block and a simplified neural network layer, and WTConv was replaced with convolutional layers of different kernel sizes and dilation rates. After the replacements, the model was retrained, and its performance was compared with that of the original model using metrics such as accuracy, precision, recall, and F1-score. The impact of these replacements on the overall performance of the model was analyzed to explore whether the performance was improved or degraded.

3) Metric: We use area under the ROC curve (AUC) as evaluation metrics.

4) Implementation Details: The size of each input is set to 6×224×224. ResNet18 and WTConv are selected as the feature extractor. The WTConv kernel size is 5, and the wavelet level is 3. Our method is implemented with PyTorch, and the experiments are performed on an NVIDIA RTX A100 GPU. During training, we set the batch size to 8. Adam optimizer is used for optimization, and the learning rate is set to $1e-4$ with a weight decay of $1e-4$. We train 50 epochs for each task.

B. Experimental results

1) Meta-Learning Ablation Experiments: The AUC value of the experimental group (using a feature extractor pre-trained based on the MAML algorithm) is 0.84, while the AUC value of the control group (training the feature extractor directly on the dataset) is 0.80. The AUC value of the experimental group is significantly higher than that of the control group, indicating that meta-learning can effectively improve the performance of the model. When faced with the task of determining the metastatic burden of breast axillary lymph nodes, we can see that the model can distinguish different situations better and improve the model's generalization ability.

2) Uni-Modal Experiments: The AUC values of models trained with only one modal image of breast MRI or axillary lymph node MRI are all lower than that of the multi-modal model. For example, the AUC value of the DCE modal model is 0.76, the AUC value of the DWI modal model is 0.71 and the AUC value of the T2W modal model is 0.68, but the AUC value of the multi-modal model reaches 0.84. This fully demonstrates that our MRI multi-modal can integrate the advantageous information of different modalities and significantly improve the accuracy of the model in determining the metastatic burden of breast axillary lymph nodes.

3) Multi-modal Experiments: The AUC value of the model with the only ResNet18.conv1 as feature extractor is 0.79. When we replace the feature extractor by the whole ResNet18, the AUC is 0.76. In contrast, the AUC of the original multi-modal model integrating DCE, T2W, and DWI modalities is 0.84. These results clearly demonstrate that our original multi-modal model outperforms the models with alternative configurations. It has a higher ability to distinguish between different classes, which implies better performance in practical applications. This indicates that the integration of multiple modalities in our model effectively captures more discriminative information, making it a more robust and accurate solution for the given task.

TABLE I
EXPERIMENTAL DATA OF THE MODEL FOR JUDGING THE METASTATIC BURDEN OF BREAST LYMPH NODES

Model & Multimodal MRI	***AUC***
MAML+ our model	0.84
our model	0.80
ResNet18.conv only	0.79
DCE	0.76
DWI	0.71
T2W	0.68

V. CONCLUSIONS

The model based on multimodal neural networks and meta-learning proposed in this paper has achieved good performance in the task of determining the metastatic burden of axillary lymph nodes in the breast. Through pre-training the feature extractor and fusing it with multimodal MRI via meta-learning, our model can fully explore the information of breast MRIs and axillary lymph node MRIs, and accurately determine the metastatic burden of axillary lymph nodes. The results of the meta-learning ablation experiment and the uni-modal experiment clearly reveal the role of meta-learning and multimodal MRI data in enhancing the model's performance and would face performance dropoffs when components of our proposed model were taken off. This proposed model has valuable clinical applications and can provide effective support for the diagnosis and treatments (both invasive and non-invasive) of breast cancer.

VI. ACKNOWLEDGMENT

This research is supported by Hong Kong ITC MHKJFS grant MHP/054/22.

REFERENCES

[1] Nolan, E., Lindeman, G.J. and Visvader, J.E., 2023. Deciphering breast cancer: from biology to the clinic. Cell, 186(8), pp.1708-1728.

[2] Giammarile, F., Vidal-Sicart, S., Paez, D., Pellet, O., Enrique, E.L., Mikhail-Lette, M., Morozova, O., Camila, N.M.M., Ivonne, R.S.D., Bolton, R.C.D. and Olmos, R.A.V., 2022. Sentinel lymph node methods in breast cancer. In Seminars in nuclear medicine (Vol. 52, No. 5, pp. 551-560). WB Saunders.

[3] Giuliano, A.E., McCall, L. and Beitsch, P., 2010. A randomized trial of axillary node dissection in women with clinical T1-2 N0 M0 breast cancer who have a positive sentinel node. the American College of Surgeons Oncology Group Z0011 randomized trial. J Clin Oncol. https://doi.org/10.1200/jco.

[4] Morrow, M., Waters, J. and Morris, E., 2011. MRI for breast cancer screening, diagnosis, and treatment. The Lancet, 378(9805), pp.1804-1811.

[5] Taghipour, M., Ziaei, A., Alessandrino, F., Hassanzadeh, E., Harisinghani, M., Vangel, M., Tempany, C.M. and Fennessy, F.M., 2019. Investigating the role of DCE-MRI, over T2 and DWI, in accurate PI-RADS v2 assessment of clinically significant peripheral zone prostate lesions as defined at radical prostatectomy. Abdominal radiology, 44, pp.1520-1527.

[6] Finn, C., Abbeel, P. and Levine, S., 2017. Model-agnostic meta-learning for fast adaptation of deep networks. In International conference on machine learning (pp. 1126-1135). PMLR.

[7] Ba, Z.C., Zhang, H.X., Liu, A.Y., Zhou, X.X., Liu, L., Wang, X.Y., Nanding, A., Sang, X.Q. and Kuai, Z.X., 2024. Combination of DCE-MRI and NME-DWI via deep neural network for predicting breast cancer molecular subtypes. Clinical Breast Cancer, 24(5), pp.e417-e427.

[8] He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[9] Finder, S.E., Amoyal, R., Treister, E. and Freifeld, O., 2024. Wavelet convolutions for large receptive fields. In European Conference on Computer Vision (pp. 363-380). Cham: Springer Nature Switzerland.

[10] Snell, J., Swersky, K. and Zemel, R., 2017. Prototypical networks for few-shot learning. Advances in neural information processing systems, 30.

[11] Vinyals, O., Blundell, C., Lillicrap, T. and Wierstra, D., 2016. Matching networks for one shot learning. Advances in neural information processing systems, 29.

[12] Lake, B., Salakhutdinov, R., Gross, J. and Tenenbaum, J., 2011. One shot learning of simple visual concepts. In Proceedings of the annual meeting of the cognitive science society (Vol. 33, No. 33).

[13] Andrychowicz, M., Denil, M., Gomez, S., Hoffman, M.W., Pfau, D., Schaul, T., Shillingford, B. and De Freitas, N., 2016. Learning to learn by gradient descent by gradient descent. Advances in neural information processing systems, 29.

[14] Finn, C., Xu, K. and Levine, S., 2018. Probabilistic model-agnostic meta-learning. Advances in neural information processing systems, 31.

[15] Zhang, P., Li, J., Wang, Y. and Pan, J., 2021. Domain adaptation for medical image segmentation: a meta-learning method. Journal of Imaging, 7(2), p.31.

[16] Mahajan, K., Sharma, M. and Vig, L., 2020. Meta-dermdiagnosis: Few-shot skin disease identification using meta-learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 730-731).

[17] Hu, Y., Zhong, Z., Wang, R., Liu, H., Tan, Z. and Zheng, W.S., 2021. Data augmentation in logit space for medical image classification with limited training data. In Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part V 24 (pp. 469-479). Springer International Publishing.

[18] Ding, X., Liu, Y., Zhao, J., Wang, R., Li, C., Luo, Q. and Shen, C., 2023. A novel wavelet-transform-based convolution classification network for cervical lymph node metastasis of papillary thyroid carcinoma in ultrasound images. Computerized Medical Imaging and Graphics, 109, p.102298.

[19] Xin, Q., Hu, S., Liu, S., Zhao, L. and Zhang, Y.D., 2022. An attention-based wavelet convolution neural network for epilepsy EEG classification. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 30, pp.957-966.

[20] Tsai, Y.H.H., Bai, S., Liang, P.P., Kolter, J.Z., Morency, L.P. and Salakhutdinov, R., 2019. Multimodal transformer for unaligned multimodal language sequences. In Proceedings of the conference. Association for computational linguistics. Meeting (Vol. 2019, p. 6558).