Multimodal Fusion Network with Distribution-based Tumor-Marker Imputation for Multi-Origin Metastatic Cervical Lymphadenopathy Classification

Rui Li[1†], Chunyan Li[2†], Xi Lin[2(✉)], Jinfeng Xu[3], Fang Li[4], and Yao Lu[1(✉)]

[1] School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, 510275, China
luyao23@mail.sysu.edu.cn

[2] Department of Ultrasound, State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, Guangzhou, 510060, China
linxi@sysucc.org.cn

[3] Department of Biostatistics, City University of Hong Kong, 999077, Hong Kong

[4] Department of Ultrasound, Chongqing Key Laboratory for Intelligent Oncology in Breast Cancer, Chongqing University Cancer Hospital, Chongqing, 400030, China

Abstract. Accurate identification the primary tumor of metastatic cervical lymphadenopathy (CLA) is crucial for guiding clinical treatment, yet clinical diagnosis remains challenging due to the complexity of tracing multi-potential origins using ultrasound images and incomplete clinical information. Existing deep learning methods typically utilize the imaging semantic features from B-mode ultrasound (BUS) and color Doppler flow imaging (CDFI), or incorporate basic clinical information, neglecting the importance of patient-specific features such as tumor markers (TMs) in clinical diagnosis. To address these limitations, we propose a new multimodal imaging-features and distribution-based tumor-marker fusion network (MDFN) for five categories of CLA metastatic origins (thyroid, head and neck, respiratory, female reproductive, and digestive). First, a distribution-based TM imputation method is proposed to reconstruct missing TMs, which treats the available clinical information of each patient (such as sex, age, neck region, etc.) as a vector to construct data distributions between TMs and avoid the data bias issues. Building on these personalized TMs, we propose the first population-personalized fusion framework, which integrates semantic features related to lymph node morphology from BUS images, semantic features related to vascular distribution from CDFI images, and TM features consistent with individualized patient data, thereby simulating clinical reasoning patterns. The effectiveness of the proposed MDFN method was evaluated using extensive experimental results from 3,100 multi-origin metastatic CLA cases, achieving an area under the receiver operating characteristic (AUC) of 0.891, with corresponding accuracy, sensitivity, specificity, and

[†]Equal contribution; ✉ Corresponding authors

F_1 of 0.863, 0.604, 0.913, and 0.661, respectively, outperforming other state-of-the-art methods.

Keywords: Metastatic cervical lymphadenopathy · Multimodal imaging-features · Distribution-based tumor-marker.

1 Introduction

Metastatic cervical lymphadenopathy (CLA), a systemic manifestation of malignancy, represents one of the most prevalent secondary cancers in patients over 40 years of age [12]. It originates from diverse primary tumors, including thyroid carcinoma, head and neck neoplasms, nasopharyngeal cancer, and pulmonary malignancies, etc., whose accurate identification becomes crucial for formulating appropriate therapeutic strategies [2]. Current clinical protocols typically require extensive diagnostic workups to localize the primary tumor sites of metastatic CLA [9]. This underscores the imperative need to develop effective, non-invasive computer-aided diagnosis (CAD) methods for identifying primary tumor sites.

Current non-invasive diagnostic approaches predominantly utilize conventional ultrasound modalities, encompassing B-mode ultrasound (BUS) and color Doppler flow imaging (CDFI). Radiologists primarily pay more attention to tumor morphology and anatomical features based on BUS and vascular structure features based on CDFI [3]. Nevertheless, the inherent pathological heterogeneity of metastatic CLA poses significant challenges in determining primary tumor origins through conventional imaging alone. Existing research predominantly focuses on distinguishing benign from malignant lymphadenopathy [7] or detecting metastatic involvement [16], with only one published study attempting metastatic CLA classification using BUS [19]. This preliminary investigation achieved an AUC of 0.822 through multi-modality integration of BUS, CDFI, ultrasound elastography (UE), contrast-enhanced ultrasound (CEUS), and clinical parameters in a limited cohort (n=280). However, the clinical applicability of the study was hampered by the small sample size and the omission of mapping the relationships between patient-specific characteristics, such as tumor markers (TMs) [17], which are critical for clinical diagnosis.

To better align with clinical diagnosis, we first proposed a new multimodal fusion network, which integrates distribution-based TMs with morphological features from BUS and vascular patterns from CDFI, enabling the classification of five-category primary tumor origins (head and neck, thyroid, respiratory system, female reproduction system, and digestive system), including 90% metastatic CLA [6]. Our contributions are summarized below:

(1) Distribution-based TM imputation: A new imputation method introduces the idea of adversarial learning to reconstruct missing TM data from each patient's clinical information, explicitly maintaining the multivariate correlation structures among TMs through cycle-consistency regularization.

(2) Multimodal imaging features and distribution-based TM fusion: A new multimodal and distribution-based fusion method is proposed, which integrates

anatomically related features extracted from BUS images, hemodynamic related features extracted from CDFI images, and distribution-based TMs profiles after imputation, thereby simulating the clinical diagnosis workflow.

(3) Large-scale multi-origin metastatic classification: Establishment of the first multi-modality population-based features with personalized TMs for five-category primary origins determination in 3,100 histologically confirmed cases, significantly outperforming previous attempts in both scope and accuracy.

2 Related work

Ultrasound-based Related work: Due to the relevant literature on multi-origin CLA prediction, this mainly introduces the classification of medical lesions similar to it. Notably, recent research [18] [10] demonstrates benign-malignant classification accuracy rates of 85%-90% when integrating multi-modality ultrasound features (incorporating BUS, CDFI, and CEUS, etc.). Although convolutional neural network architecture has shown effectiveness in deep semantic feature extraction, its feature space remains population-based, and the integration of the mapping relationship between each patient's features is still an area to be explored.

Data imputation-based related work: Missing clinical parameters represent a pervasive challenge in medical artificial intelligence, where accurate data imputation critically determines diagnostic model performance. Current imputation methodologies bifurcate into statistical approaches (e.g., Multiple Imputation by Chained Equations [14], matrix decomposition [11]) and deep learning methods. The former leverage mathematical optimization to preserve feature correlations, whereas the latter construct nonlinear mapping spaces through generative architectures like adversarial networks[5]. However, the essence of the above research is still based on a population-based image feature space, overlooking the interaction between different clinical data of the same patient.

3 Method

As shown in Fig.1, the proposed multimodal imaging-features and distribution-based tumor-marker fusion network (MDFN) [5] primarily comprises BUS branch, CDFI branch, distribution-based TMs imputation branch, and feature fusion branch.

3.1 BUS Branch

For clinicians focusing on morphological information in BUS images, this branch employs the InceptionResNetV2 architecture [13] to extract population-based, morphology-based semantic feature. This multiscale convolutional backbone synergistically integrates Inception-style parallel pathways (with 1×1, 3×3, and

[5] The code can be found at https://github.com/lr256/MDFN.

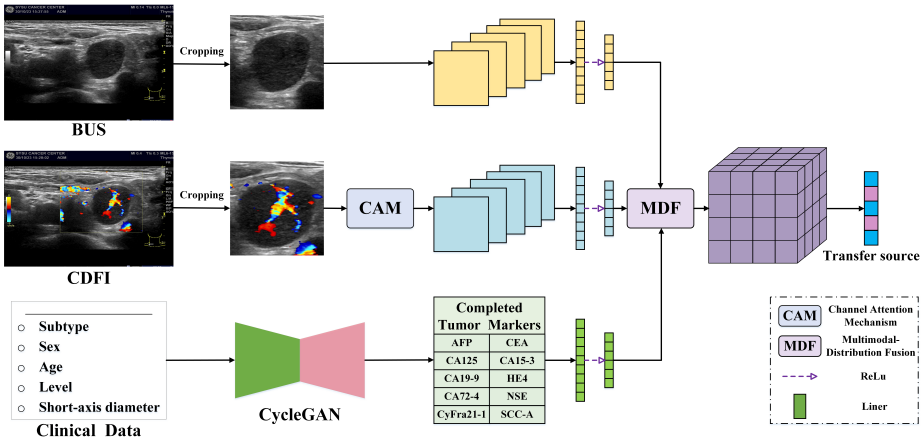

Fig. 1. Proposed framework of the multimodal imaging-features and distribution-based tumor-marker fusion network (MDFN). BUS images were analyzed to extract anatomical semantic features, while CDFI images were analyzed to extract vascular distribution features indicated by colorful dots

5×5 convolutional filters) and residual skip connections, enabling hierarchical learning of discriminative morphological patterns while mitigating gradient vanishing in deep networks. The architecture's cross-channel normalization and batch-aware feature recalibration mechanisms enhance clinically relevant feature discovery, particularly for subtle textural variations critical in BUS interpretation. Following depthwise feature aggregation through a nonlinear projection module, we apply dimensionality reduction via a learnable projection layer (1536-d to 1024-d), generating fixed-dimensional semantic embeddings.

3.2 CDFI Branch

The CDFI branch employs a two-stage attention guidance architecture to extract population-based semantic features related to vascular patterns. First, raw Doppler signals amplify hemodynamically salient color components through a 3-channel (RGB) Channel Attention Mechanism module while suppressing acquisition artifacts [8]. This channel-adaptive feature recalibration precedes Inception-ResNetV2 processing to preserve directional flow signatures critical for microvascular assessment. The backbone network then extracts directional hemodynamic features through its multiscale receptive fields, with particular emphasis on vessel tortuosity and branching patterns via depthwise separable convolutions. Following feature aggregation, a nonlinear projection module (two fully-connected layers with ReLU) reduces dimensionality from 1536-d to 1024-d.

3.3 Distribution-based TMs Imputation Branch

To address the pervasive challenge of missing TMs in clinical oncology and maintain the mapping between patient-specific features, we developed a conditional adversarial imputation framework grounded in domain-specific clinical prior knowledge. Our CycleGAN architecture [4] establishes personalized bidirectional clinical-feature-to-marker mappings using five key diagnostic determinants: tumor subtype, age, sex, lesion level, and short-axis diameter, thereby emulating clinical diagnostic workflows. Specifically, each of the 10 tumor markers (AFP, CEA, CA125, CA15-3, CA19-9, HE4, CA72-4, NSE, CyFra21-1, SCC-A) employs a dedicated CycleGAN unit comprising:Generator (G): 4-layer MLP [1] with residual connections (ReLU) translating clinical parameters to marker values. (D): Wasserstein GAN critic with gradient penalty ($\lambda = 10$) enforcing distributional consistency between real/generated markers. In the same way, we also map it to a unified 1024-d using a nonlinear projection module.

3.4 Multimodal-Distribution Fusion Branch

To achieve an effective fusion of multimodal and distribution-based features (morphological-related semantic features extracted from BUS, hemodynamic features extracted from CDFI, and personalized TM features), we employ hierarchical tensor fusion network [15]. It constructs a joint representation space of cross-modal feature interactions by calculating the outer product of three feature vectors, capturing nonlinear and higher-order correlation between different modes. Finally, the low-dimensional features, after tensor decomposition, are mapped to the probability of primary tumor origin of CLA metastasis through a linear connection layer.

4 Results and Discussion

This study retrospectively analyzed the data of 3100 metastatic cervical lymph nodes from Sun Yat-sen University Cancer Center, of which 280 were from thyroid, 218 were from head and neck, 1249 were from respiratory system, 910 were from female reproductive system, and 443 were from digestive system. The missing ratios for these TMs were 54.5%, 3.6%, 57.6%, 39.7%, 36.9%, 84.9%, 60.1%, 36.1%, 53.3%, and 85.9%, respectively. We divided the dataset into training and test sets with an 8:2 ratio and performed 3-fold cross-validation on the training. We conducted comparative experiments to evaluate the method's effectiveness.

4.1 Analysis of Classification Performance

The experimental results validate the classification performance of the proposed MDFN framework compared with different modalities, as shown in Table 1. Comparative analysis of two single-modality methods revealed a 3.9% improvement in AUC for M_{BUS} over M_{CDFI} (0.743 vs 0.704), aligning with radiologists'

predominant reliance on morphological features in clinical practice. Although the AUC of $M_{BUS+CDFI}$ increased by 0.3% (0.746 vs 0.743), the F_1 was significantly improved by 10.4% (0.445 vs 0.341), indicating enhanced precision-recall balance. Crucially, integration of imputed TM with imaging data yielded significant performance gains across all modality. The AUC of baseline M_{BUS} improved by 9.7% when combined with TM (M_{BUS+TM}). Similarly, CDFI combined with TM ($M_{CDFI+TM}$) elevated AUC from 0.704 to 0.806 (increased by 10.2%), validating the complementary diagnostic value of TMs. Compared with other methods, the proposed method MDFN effectively fused the heterogeneous characteristics of multimodal imaging and distribution-based TMs, achieving the optimal value among five evaluation indicators (AUC, Accuracy, Sensitivity, Specificity, F_1): AUC and F_1 are 5.1% and 11.1% higher, respectively, than the next best method (M_{BUS+TM}). Fig. 2 (a) depicts the receiver operating characteristic curve of the corresponding AUC value, where proposed MDFN achieves an optimal AUC of 89.1%, surpassing comparator methods across all specificity thresholds. These findings collectively demonstrate MDFN's capacity to leverage cross-modal synergies through tensor fusion, greatly improving classification accuracy for metastatic cervical lymphadenopathy.

Table 1. The comparison of classification performance under different modality data.

Method	AUC (95% CI)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	F_1
M_{BUS}	0.743 (0.718-0.765)	0.792 (0.777-0.809)	0.413 (0.383-0.443)	0.861 (0.851-0.872)	0.341
M_{CDFI}	0.704 (0.676-0.729)	0.751 (0.736-0.767)	0.293 (0.261-0.324)	0.823 (0.813-0.834)	0.244
$M_{BUS+CDFI}$	0.746 (0.721-0.768)	0.786 (0.771-0.802)	0.399 (0.369-0.432)	0.858 (0.848-0.868)	0.445
M_{BUS+TM}	0.840 (0.816-0.862)	0.848 (0.833-0.862)	0.542 (0.495-0.59)	0.901 (0.891-0.910)	0.550
$M_{CDFI+TM}$	0.806 (0.778-0.831)	0.823 (0.809-0.839)	0.481 (0.439-0.528)	0.880 (0.871-0.891)	0.489
$MDFN$	0.891 (0.872-0.908)	0.863 (0.847-0.878)	0.604 (0.561-0.647)	0.913 (0.904-0.923)	0.661

4.2 Comparison with Different Imputation Method

The CycleGAN-based tumor marker imputation module constitutes a core module of proposed MDFN method, enabling robust prediction of multi-origin metastases. Fig. 2 (b) and Table 2 provides the ROC and a quantitative comparison of the different imputation methods used in predicting multiple transfer origins.

Table 2. The performance comparison under different imputation methods.

Method	AUC (95% CI)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	F_1
MICE	0.858 (0.837-0.878)	0.860 (0.845-0.874)	0.558 (0.515-0.604)	0.906 (0.897-0.915)	0.586
Average padding	0.866 (0.846-0.883)	0.825 (0.809-0.842)	0.523 (0.481-0.568)	0.886 (0.876-0.896)	0.491
Zero padding	0.870 (0.849-0.891)	0.845 (0.831-0.861)	0.592 (0.547-0.64)	0.900 (0.891-0.911)	0.585
Logistic regression	0.871 (0.850-0.890)	0.851 (0.837-0.865)	0.593 (0.544-0.637)	0.904 (0.895-0.913)	0.588
Our-CycleGAN	0.891 (0.872-0.908)	0.863 (0.847-0.878)	0.604 (0.561-0.647)	0.913 (0.904-0.923)	0.661

Compared to the three traditional imputation methods, fixed-value padding (average padding and zero pading) yields better classification performance (0.886 vs 0.870 vs 0.858) than the Multiple Imputation by Chained Equations (MICE) method. This may be attributed to their conservative preservation of original TM distributions rather than introducing spurious correlations through statistical assumptions. In contrast, the proposed personalized imputation method based on each patient's clinical information effectively simulates the complex relationship between TMs and achieves the best classification performance among all evaluation indicators, outperforming other methods.

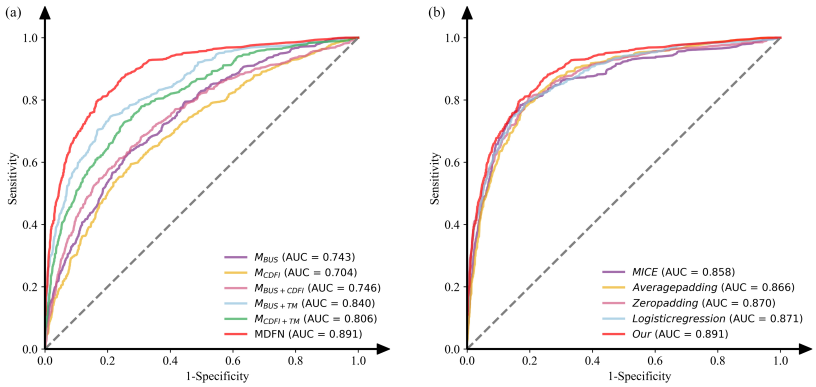

Fig. 2. ROCs used to evaluate the performance of different methods.

Table 3. Performance of the model in ablation study

Method	AUC (95% CI)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	F_1
MDFN (None CAM)	0.876 (0.858-0.894)	0.849 (0.834-0.864)	0.573 (0.535-0.613)	0.904 (0.895-0.914)	0.536
MDFN (None MDF)	0.874 (0.855-0.891)	0.847 (0.832-0.862)	0.567 (0.526-0.608)	0.902 (0.892-0.912)	0.530

4.3 Ablation Study

Building on the comprehensive ablation studies in Section 4.1 (which compare different modalities and the presence/absence of tumor markers), this analysis focuses on the two remaining innovative components: 1) the Channel Attention Mechanism (CAM) for context-aware feature recalibration, and 2) Multimodal-Distribution Fusion (MDF) for cross-modal interaction modeling. As empirically demonstrated in Table 3, both modules showed consistent performance improvements: the CAM implementation increased the AUC from 0.876 to 0.891 ($\triangle$+1.5%), enhancing sensitivity to vascular pattern variations through adaptive channel weighting. The addition of TF module raised the AUC from 0.874 to 0.891 and increased the F_1 by 13.1%, confirming its effectiveness in fusion imgaing and distribution-based features.

Table 4. The comparison of state-of-the art method

Reference	Method	AUC (95% CI)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	F_1
Zhu et al., 2024 [19]	MSMFN	0.633 (0.605-0.658)	0.794 (0.779-0.809)	0.290 (0.274-0.304)	0.847 (0.838-0.857)	0.229
Our	MDFN	0.891 (0.872-0.908)	0.863 (0.847-0.878)	0.604 (0.561-0.647)	0.913 (0.904-0.923)	0.661

4.4 Comparison with Previous State-of-the-art Method

Given the limited prior research on ultrasound-based multi-origin metastasis prediction in cervical lymphadenopathy (CLA), we conducted a comparative analysis with the only existing method (MSMFN) presented in Table 4. Although MSMFN originally incorporates UE and CEUS modalities alongside BUS/CDFI, we strategically disabled these components in our implementation to ensure a fair comparison based solely on core ultrasound modalities. In comparison, the

proposed MDFN method aligns more closely with the diagnostic approach due to the integration of distribution-based completed TMs and dual-modality semantic features, thus improving the classification performance.

5 Conclusion

We proposed a population-personalized fusion network for identifying five-category metastatic CLA. Evaluating its effectiveness on a dataset of 3,100 malignant CLA cases from multiple metastatic origins, the method outperforms single-modality methods and those without TM imputation, achieving an AUC of 0.891, accuracy of 0.863, sensitivity of 0.604, specificity of 0.913, and F_1 score of 0.661. Additionally, the distribution-based TM imputation method from existing clinical data to preserve interaction relationships and integrates dual-modality image features, to improve multi-class classification performance.

Acknowledgments. This work was funded in part by the National Key R&D Program of China under Grant 2023YFE0204300; in part by the R&D project of Pazhou Lab (HuangPu) under Grant 2023K0606; in part by the National Natural Science Foundation of China under Grants 82441027, 62371476, and 82171955; in part by the Guangzhou Science and Technology Bureau under Grant 2023B03J1237.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Abbas, D.K., Rashid, T.A., Bacanin, K.H., Alsadoon, A.: Using fitness dependent optimizer for training multi-layer perceptron. arXiv preprint arXiv:2201.00563 (2022)
2. Alečković, M., McAllister, S.S., Polyak, K.: Metastasis as a systemic disease: molecular insights and clinical implications. Biochimica et Biophysica Acta (BBA)-Reviews on Cancer **1872**(1), 89–102 (2019)
3. Cheung, Y.C., Wan, Y.L., Lui, K.W., Lee, K.F.: Sonographically guided core-needle biopsy in the diagnosis of superficial lymphadenopathy. Journal of clinical ultrasound **28**(6), 283–289 (2000)
4. Chu, C., Zhmoginov, A., Sandler, M.: Cyclegan, a master of steganography. arXiv preprint arXiv:1712.02950 (2017)
5. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial networks. Communications of the ACM **63**(11), 139–144 (2020)
6. Han, F., Xu, M., Xie, T., Wang, J.W., Lin, Q.G., Guo, Z.X., Zheng, W., Han, J., Lin, X., Zou, R.H., et al.: Efficacy of ultrasound-guided core needle biopsy in cervical lymphadenopathy: A retrospective study of 6,695 cases. European radiology **28**, 1809–1817 (2018)
7. Lee, J.H., Baek, J.H., Kim, J.H., Shim, W.H., Chung, S.R., Choi, Y.J., Lee, J.H.: Deep learning–based computer-aided diagnosis system for localization and diagnosis of metastatic lymph nodes on ultrasound: a pilot study. Thyroid **28**(10), 1332–1338 (2018)

8. Li, H., Qiu, K., Chen, L., Mei, X., Hong, L., Tao, C.: Scattnet: Semantic segmentation network with spatial and channel attention mechanism for high-resolution remote sensing images. IEEE Geoscience and Remote Sensing Letters **18**(5), 905–909 (2020)
9. Li, Y., Zhang, Q., Xiang, T., Lin, Y., Zhang, Q., Li, X.: Few-Shot Lymph Node Metastasis Classification Meets High Performance on Whole Slide Images via the Informative Non-Parametric Classifier . In: proceedings of Medical Image Computing and Computer Assisted Intervention – MICCAI 2024. vol. LNCS 15012. Springer Nature Switzerland (October 2024)
10. Liu, Y., Chen, J., Zhang, C., Li, Q., Zhou, H., Zeng, Y., Zhang, Y., Li, J., Xv, W., Li, W., et al.: Ultrasound-based radiomics can classify the etiology of cervical lymphadenopathy: A multi-center retrospective study. Frontiers in Oncology **12**, 856605 (2022)
11. Ni, Z., Zheng, X., Zheng, X., Zou, X.: sclrtd: A novel low rank tensor decomposition method for imputing missing values in single-cell multi-omics sequencing data. IEEE/ACM Transactions on Computational Biology and Bioinformatics **19**(2), 1144–1153 (2020)
12. Pynnonen, M.A., Gillespie, M.B., Roman, B., Rosenfeld, R.M., Tunkel, D.E., Bontempo, L., Brook, I., Chick, D.A., Colandrea, M., Finestone, S.A., et al.: Clinical practice guideline: evaluation of the neck mass in adults. Otolaryngology–Head and Neck Surgery **157**(2_suppl), S1–S30 (2017)
13. Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.: Inception-v4, inception-resnet and the impact of residual connections on learning. In: Proceedings of the AAAI conference on artificial intelligence. vol. 31 (2017)
14. Van Buuren, S., Groothuis-Oudshoorn, K.: mice: Multivariate imputation by chained equations in r. Journal of statistical software **45**, 1–67 (2011)
15. Wang, T., Chen, H., Chen, Z., Li, M., Lu, Y.: Prediction model of early recurrence of multimodal hepatocellular carcinoma with tensor fusion. Physics in Medicine & Biology **69**(12), 125003 (2024)
16. You, J., Huang, Y., Ouyang, L., Zhang, X., Chen, P., Wu, X., Jin, Z., Shen, H., Zhang, L., Chen, Q., et al.: Automated and reusable deep learning (autordl) framework for predicting response to neoadjuvant chemotherapy and axillary lymph node metastasis in breast cancer using ultrasound images: a retrospective, multicentre study. EClinicalMedicine **69** (2024)
17. Yuan, C., Wang, Y., Li, Y.: Tmdfm: A data fusion model for combined detection of tumor markers. In: 2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). pp. 657–660. IEEE (2015)
18. Zhu, Y., Meng, Z., Fan, X., Duan, Y., Jia, Y., Dong, T., Wang, Y., Song, J., Tian, J., Wang, K., et al.: Deep learning radiomics of dual-modality ultrasound images for hierarchical diagnosis of unexplained cervical lymphadenopathy. BMC medicine **20**(1), 269 (2022)
19. Zhu, Y., Meng, Z., Wu, H., Fan, X., Tian, J., Wang, K., Nie, F., et al.: Deep learning radiomics of multimodal ultrasound for classifying metastatic cervical lymphadenopathy into primary cancer sites: a feasibility study. Ultraschall in der Medizin-European Journal of Ultrasound **45**(03), 305–315 (2024)