Domain-Bridged Visual Adaptation Framework for Clinical Ultrasound Image Understanding

Se-woon Choe
Department of Biomedical Engineering
Kumoh National Institute of Technology
Gumi, Republic of Korea
sewoon@kumoh.ac.kr

Abstract— **This study presents a domain-bridged visual adaptation framework for improving automated ultrasound image classification with limited labeled data. The approach employs a multistage transfer-learning pipeline that progressively aligns feature representations across natural images, intermediate medical imagery, and ultrasound-specific patterns. This gradual transition fosters stable and clinically meaningful feature learning while reducing overfitting. The paper outlines the conceptual motivation, dataset design, training strategy, and interpretability analyses. Experimental results show that the proposed framework enhances classification accuracy, training stability, and cross-dataset generalization, demonstrating its effectiveness for robust ultrasound analysis.**

Keywords— **Domain bridge learning, cross-domain adaptation, ultrasound imaging, transfer learning, medical AI**

I. INTRODUCTION

Ultrasound imaging plays an essential role in contemporary medical diagnostics due to its affordability, real-time acquisition capability, and lack of ionizing radiation [1]. These advantages have led to its widespread adoption in applications such as breast cancer screening, abdominal examinations, obstetric assessments, and cardiovascular analysis [2]. Despite its clinical value, automated interpretation of ultrasound images remains difficult. The modality inherently suffers from speckle noise, acoustic shadowing, blurred boundaries, and operator-dependent variations in probe orientation, all of which introduce substantial inconsistencies in appearance [3]. These factors complicate the extraction of coherent feature representations and limit the effectiveness of deep neural networks trained on modestly sized ultrasound datasets [4].

The scarcity of large, well-annotated ultrasound datasets further exacerbates these challenges. Unlike natural-image collections, which often contain millions of labeled examples, publicly available ultrasound datasets are comparatively small and heterogeneous [5]. Models trained from scratch on such limited data frequently exhibit unstable convergence, high sensitivity to noise, and a strong tendency to overfit [6].

Transfer learning offers a partial remedy by leveraging feature representations learned from natural-image datasets such as ImageNet [7]. Although pretrained models provide robust low-level filters and general visual descriptors, their internal representations do not naturally align with the diffuse textures and ambiguous structures characteristic of ultrasound imagery [8]. Natural photographs contain well-defined object boundaries, rich color gradients, and clear geometric features, whereas ultrasound images display irregular textures, variable intensity patterns, and low-contrast lesion margins. As a result, directly fine-tuning natural-image models typically yields inconsistent and suboptimal performance [9].

Recent studies have explored domain adaptation strategies to reduce this discrepancy. However, many existing methods rely on adversarial objectives or complex distribution-alignment losses and often require extensive intermediate datasets that may not be readily accessible [10]. Furthermore, directly adapting a model from natural images to ultrasound imagery involves a considerable representational leap due to their fundamentally different visual characteristics [11].

To address these limitations, this study introduces a Domain-Bridged Cross-Adaptation Framework that guides the model through a smoother representational transition. Instead of applying a single adaptation step, the framework progressively shifts the model's feature space from natural-image representations to intermediate medical-image structures and finally to ultrasound-specific patterns. This staged progression encourages the formation of medically aligned features, reduces the burden of abrupt domain shifts, improves stability during training, and enhances sensitivity to fine-grained lesion characteristics. The following sections describe the conceptual foundations, design strategy, and empirical validation of the proposed approach.

II. MATERIALS AND METHODS

Deep neural networks develop increasingly abstract visual representations across their hierarchical layers. Early layers capture edges, gradients, and low-frequency textures, while deeper layers encode shape-based and semantic information. Transfer learning assumes that these representations retain some degree of utility across tasks, yet the semantic gap between natural images and ultrasound images is large enough to limit their direct transferability [7].

To mitigate this discrepancy, the proposed domain-bridged strategy introduces an intermediate medical-image domain that shares structural similarities with ultrasound imagery [11]. Medical images produced by modalities such as microscopy, radiography, and dermoscopy often contain irregular boundaries, heterogeneous textures, and locally varying

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education and the Ministry of Science and ICT under grant numbers RS-2023-00240521 and RS-2025-16068429.

intensity profiles—properties that are more aligned with the characteristics of ultrasound images than with those of natural photographs. Training the model on such imagery creates a representational staging ground that encourages medically meaningful feature extraction before the final adaptation to ultrasound data.

The first stage initializes the model using weights learned from a large-scale dataset of natural images [7]. This initial phase provides a stable and expressive set of filters capable of capturing diverse low-level features. Although these representations do not reflect medical structures, they offer a reliable foundation that prevents unstable learning behaviors when the model encounters small medical datasets in subsequent stages. The pretrained backbone also accelerates convergence and reduces the likelihood of vanishing gradients during early training iterations.

In the second stage, the entire network is retrained using an intermediate dataset composed of structurally rich medical images [11]. These images may include microscopic cellular structures, dermoscopic skin patterns, or radiographic textures, all of which exhibit visual characteristics that resemble those observed in ultrasound scans. Retraining on this dataset encourages the model to depart from the object-centric semantics prevalent in natural images and instead adopt more medically oriented abstractions. This stage enables the development of intermediate representations that bridge the disparity between natural-image features and ultrasound-specific patterns, thereby reducing the learning burden during the final fine-tuning phase [12].

In the final stage, the model undergoes fine-tuning on ultrasound datasets containing benign and malignant breast lesions [5]. All layers remain trainable so that the feature representations learned in earlier stages can be fully realigned with the unique textural and structural properties of ultrasound imaging. Minimal preprocessing is applied to maintain generalizability; images are resized, normalized, and augmented with realistic transformations such as moderate rotation, flipping, contrast adjustment, and intensity perturbation. These augmentations mimic variability encountered in clinical practice. The Adam optimizer, paired with a cosine decay learning-rate schedule, is used to maintain stable convergence. Light regularization with dropout and batch normalization is employed to stabilize feature distributions while preserving clinically relevant textures [13].

The framework is evaluated using multiple backbone architectures and consistent training configurations to ensure a fair comparison. Performance is assessed via accuracy, sensitivity, specificity, F1-score, and AUC [14]. Cross-dataset experiments are conducted to examine robustness to domain shifts, while ablation studies quantify the specific contribution of the intermediate medical-domain adaptation. Qualitative analyses involving activation maps and feature visualizations provide insight into the internal representational changes induced by the domain-bridged training process.

III. Results

Experiments demonstrate that the domain-bridged framework consistently outperforms conventional transfer learning. Across various CNN backbones, models trained using the proposed multistage strategy achieve higher accuracy, improved sensitivity to malignant cases, and reduced variance across validation folds. Residual architectures in particular yield the strongest performance, likely due to their ability to preserve gradient flow and refine deeper representational features during staged adaptation [15]. In contrast, EfficientNet and Inception-based models show moderate but less stable improvements.

Single-stage transfer learning often results in unstable convergence, characterized by oscillations in loss and accuracy. These instabilities stem from the abrupt transition between natural-image representations and ultrasound imagery. After incorporating the intermediate medical-domain stage, the training curves become markedly smoother and more consistent. The model demonstrates a reduced tendency to memorize noise patterns and instead learns more generalizable textural distinctions. This stability is maintained even in datasets with limited sample sizes.

The multistage framework yields classification boundaries that more accurately separate benign and malignant lesions. ROC analyses reveal that the proposed method produces curves with stronger class separation and higher AUC scores. These trends indicate improved discriminative capacity and more reliable clinical decision support. Statistical evaluations confirm that the observed improvements are significant and unlikely to result from random variation.

Testing on an independent ultrasound dataset further validates the robustness of the domain-bridged approach. The model maintains high sensitivity and specificity despite differences in acquisition equipment, imaging quality, and patient demographics. This outcome highlights the framework's ability to generalize across diverse clinical settings and demonstrates its resilience to domain shifts that commonly challenge medical imaging models.

Visualization of learned feature maps reveals clear distinctions between the domain-bridged model and conventional transfer-learning baselines. While single-stage models tend to focus on strong edges and global patterns, the domain-bridged framework captures fine-grained textures, irregular lesion boundaries, and subtle structural cues associated with malignancy. The progressive transition established in earlier stages appears to reorganize feature channels into medically meaningful clusters, indicating the emergence of domain-specific abstractions.

IV. Conclusion

This work presents a structured domain-bridged learning framework that progressively transitions visual representations from natural images through intermediate medical imagery and ultimately to ultrasound data. By incorporating a staged adaptation process, the framework reduces the representational mismatch that hampers conventional transfer learning and enables the extraction of stable, clinically relevant features from limited ultrasound datasets. Extensive evaluation demonstrates that the method enhances accuracy, sensitivity, interpretability, and cross-dataset robustness.

Although this study focuses on breast ultrasound classification, the proposed approach is broadly applicable to other medical imaging tasks where annotated data are limited. Future research may explore alternative intermediate domains, multimodal pretraining sources, larger multi-center datasets, and integration with self-supervised learning strategies. Overall, the domain-bridged framework offers a highly promising foundation for developing reliable, adaptable, and clinically deployable AI systems for ultrasound image analysis.

ACKNOWLEDGMENT

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education and the Ministry of Science and ICT under grant numbers RS-2023-00240521 and RS-2025-16068429.

REFERENCES

[1] J. Geisel, M. Raghu, and R. Hooley, "The role of ultrasound in breast cancer screening: The case for and against ultrasound," *Semin. Ultrasound CT MRI*, vol. 39, pp. 25–34, Jan. 2018.

[2] J. M. Seely and T. Alhassan, "Screening for breast cancer in 2018—What should we be doing today?" *Curr. Oncol.*, vol. 25, pp. S115–S124, Jan. 2018.

[3] S. Liu, Y. Wang, X. Yang, B. Lei, L. Liu, S. Li, D. Ni, and T. Wang, "Deep learning in medical ultrasound analysis," *Engineering*, vol. 5, pp. 261–275, Apr. 2019.

[4] G. Ayana, K. Dese, and S. Choe, "Transfer learning in breast cancer diagnoses via ultrasound imaging," *Cancers*, vol. 13, p. 738, Feb. 2021.

[5] W. Al-Dhabyani, M. Gomaa, H. Khaled, and A. Fahmy, "Dataset of breast ultrasound images," *Data Brief*, vol. 28, p. 104863, Apr. 2020.

[6] M. A. Morid, A. Borjali, and G. Del Fiol, "A scoping review of transfer learning research on medical image analysis using ImageNet," *Comput. Biol. Med.*, vol. 128, p. 104115, Mar. 2021.

[7] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, "A comprehensive survey on transfer learning," *Proc. IEEE*, vol. 109, pp. 43–76, Jan. 2021.

[8] M. Byra, M. Galperin, H. Ojeda-Fournier, L. Olson, M. O'Boyle, C. Comstock, and M. Andre, "Breast mass classification in sonography with transfer learning," *Med. Phys.*, vol. 46, pp. 746–755, Feb. 2019.

[9] N. Houssami, G. Kirkpatrick-Jones, N. Noguchi, and C. I. Lee, "Artificial intelligence for early detection of breast cancer," *Expert Rev. Med. Devices*, vol. 16, pp. 351–362, Apr. 2019.

[10] J. R. Burt, N. Torosdagli, N. Khosravan, H. RaviPrakash, A. Mortazi, F. Tissavirasingham, S. Hussein, and U. Bagci, "Recent advances in diagnosing breast cancer with deep neural networks," *Br. J. Radiol.*, vol. 91, p. 20170545, Mar. 2018.

[11] R. K. Samala, H-P. Chan, L. Hadjiiski, M. Helvie, C. Richter, and K. Cha, "Breast cancer diagnosis in digital breast tomosynthesis: Effects of training sample size on multi-stage transfer learning," *IEEE Trans. Med. Imaging*, vol. 38, pp. 686–696, Mar. 2019.

[12] A. Zhao, G. Balakrishnan, F. Durand, J. Guttag, and A. Dalca, "Data augmentation using learned transformations," in *Proc. CVPR*, pp. 8535–8545, Jun. 2019.

[13] A. Labach, H. Salehinejad, and S. Valaee, "Survey of dropout methods for deep neural networks," *arXiv*, Apr. 2019.

[14] K. Hajian-Tilaki, "Receiver operating characteristic (ROC) curve analysis," *Casp. J. Intern. Med.*, vol. 4, pp. 627–635, Dec. 2013.

[15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. CVPR*, pp. 770–778, Jun. 2016.