Cancer Biol Med 2026. doi: 10.20892/j.issn.2095-3941.2025.0747

ORIGINAL ARTICLE

Virtual histology imaging of lymph nodes *via* dynamic full-field optical coherence tomography and deep learning to differentiate metastasis

Shuwei Zhang[1]*, Houpu Yang[1]*, Yiyin Zhang[1]*, Xiaoxian Li[2], Jin Zhao[1], Yuanyuan Zhang[3], Ping Xue[4], Hua Kang[5], Hongchuan Jiang[6], Wenhui Ren[7], Shu Wang[1]

[1]Breast Center, Peking University People's Hospital, Beijing 100044, China; [2]Department of Pathology and Laboratory Medicine, Emory University, Atlanta, GA 30322, USA; [3]Department of Pathology, Peking University People's Hospital, Beijing 100044, China; [4]Department of Physics and State Key Laboratory of Low-dimensional Quantum Physics, Tsinghua University, Beijing 100084, China; [5]Department of General Surgery, Xuanwu Hospital, Capital Medical University, Beijing 100053, China; [6]Department of Breast Surgery, Beijing Chaoyang Hospital, Capital Medical University, Beijing 100020, China; [7]Department of Clinical Epidemiology and Biostatistics, Peking University People's Hospital, Beijing 100044, China

ABSTRACT

Objective: The current pathological diagnosis of lymph node metastasis is time-consuming, labor-intensive, and dependent on sectioning of paraffin blocks. Herein, in a prospective cohort of patients with breast cancer, we validated dynamic full-field optical coherence tomography (D-FFOCT), a virtual pathology tool integrating deep learning for nodal metastasis detection, and offering rapid and label-free histologic approximations of fresh tissues.

Methods: In a prospective dual-center cohort of 155 patients with breast cancer, 747 freshly bisected lymph node slides were obtained *via* D-FFOCT. Surgeons interpreted each slide with histopathology as the gold standard. A deep learning model was trained on 28,911 patches (corresponding to 590 slides) and tested on 7,736 patches (corresponding to 157 slides). The results were mapped to the slide level for potential intraoperative evaluation.

Results: D-FFOCT strongly correlated with hematoxylin and eosin (H&E)-stained histological images. Surgeons achieved 97.10% specificity in nodal diagnosis with D-FFOCT. The performance of the artificial intelligence (AI) model was not inferior to that of human experts and had a sensitivity/specificity of 87.88%/91.94% and an area under the receiver operating characteristic curve of 0.899 at the slide level. The human–AI collaborative system reduced labor requirements by 75% and increased the specificity by 6.5%, to 98.39%.

Conclusions: D-FFOCT has excellent potential as a tool for assessing lymph node metastatic status without tissue preparation or consumption. The integration of D-FFOCT with deep learning decreases labor demands and maintains high accuracy, thereby enabling streamlined nodal prediction independent of routine pathology procedures.

KEYWORDS Breast cancer; dynamic full field optical coherence tomography; lymph nodes; AI model; metastatic status

Introduction

Sentinel lymph node biopsy (SLNB) has become an important technique for the detection of regional metastasis to guide subsequent surgical procedures[1-4]. SLNB can effectively minimize impaired lymphatic reflux and decrease the risk of complications. Frozen section (FS) analysis is commonly used for intraoperative nodal assessment, but this approach remains constrained by its suboptimal diagnostic accuracy, time-intensive processing, and tissue consumption, which may compromise molecular analyses[5-7]. Paraffin sectioning-based techniques enable accurate diagnosis of nodal metastasis, but they do not offer timely diagnosis. Moreover, positive diagnoses may lead to repeat operations, and the associated hematoxylin–eosin (H&E) staining and immunohistochemistry tests may pose substantial healthcare costs.

To meet demands for improved nodal assessment procedures, various nonconsumptive and high-resolution imaging techniques have been developed as promising alternatives to real-time tissue assessment. Raman spectroscopy, although

*These authors contributed equally to this work.
Correspondence to: Shu Wang
E-mail: shuwang@pkuph.edu.cn
ORCID ID: https://orcid.org/0000-0002-8651-4795
Received November 23, 2025; accepted February 27, 2026;
published online March 28, 2026.
Available at www.cancerbiomed.org
©2026 The Authors. Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

chemically specific in providing molecular fingerprints of biological samples[8], is limited by intrinsic background interference and prolonged acquisition times[9]. Quantitative phase imaging (QPI) uses virtual staining to generate pathologist-friendly images, but it inherently relies on indirect approaches that lack ground-truth images for comparison[10]. Therefore, a noninvasive method that can deliver high-resolution histopathological information with intrinsic contrast remains needed.

Dynamic full-field optical coherence tomography (D-FFOCT), an advanced derivative of static FFOCT[11], has emerged as a label-free, fixation-free, and wide-field optical imaging modality capable of generating high-contrast, histology-like images with subcellular resolution[12]. This technique uses light interference to measure the backscattered light and gauge the cell viability of fresh tissues, thus generating contrast-enhanced virtual pathology images through pseudocoloring of spatiotemporal metabolic activity patterns[11]. D-FFOCT has recently been demonstrated to capture dynamic processes and discern cellular microstructures in *ex vivo* tissues and *in vitro* organoids[13,14]. Our pilot study in 173 breast biopsies and 141 resected lymph nodes, taking advantage of the distinct metabolic profiles of benign *vs.* malignant cells, has demonstrated the potential of D-FFOCT to differentiate benign and malignant lesions[15]. We subsequently integrated artificial intelligence with D-FFOCT for breast tumor classification and conducted a simulated margin evaluation based on large-scale D-FFOCT images, in which only small regions were selected to approximate margin status[16]. In contrast, the current study conducted an application-driven validation using D-FFOCT for intraoperative lymph node assessment.

To bridge the gap from simulation to translational applications, we systematically evaluated the clinical utility of D-FFOCT for lymph node assessment through a large-scale, prospective, dual-center cohort study (n = 747). This cohort encompassed the heterogeneity of lymph node metastases encountered in real-world clinical practice. Histopathology served as the reference standard. To minimize the labor of current intraoperative diagnostic modalities, we implemented a deep learning framework for the metastatic classification of axillary lymph nodes of patients with breast cancer. This study demonstrated the unique ability of D-FFOCT to reveal histological details with H&E-like contrast in lymph nodes without exogenous contrast agents, thus establishing an automated and pathology-independent paradigm for real-time cancer nodal assessment.

Materials and methods

Study design and participants

This dual-center prospective study (ClinicalTrials.gov NCT03791853) was designed to evaluate the diagnostic accuracy of D-FFOCT for lymph node assessment and was approved by the Peking University People's Hospital Ethics Committee (2016PHB210-001). We sequentially enrolled patients with breast cancer from 2 centers (Peking University People's Hospital Breast Center followed by Beijing Chaoyang Hospital of Capital Medical University) who required nodal staging (SLNB or axillary lymph node dissection). All participants provided written informed consent to use of their specimens and data. The participant profile of the study is summarized in **Figure 1A**.

D-FFOCT imaging

Lymph nodes were obtained from intact resected specimens during SLNB or from specimens obtained *via* axillary lymph node dissection, on the basis of gross inspection and palpation. Qualified nodes were bisected longitudinally to present a smooth, fresh tissue surface, then imaged with a D-FFOCT scanner (LLTech, Paris, France) according to established protocols[15,16]. Each sample was scanned at a depth of 10 μm to obtain a diagnostic D-FFOCT slide, then processed to generate a formalin-fixed, paraffin-embedded (FFPE) block for subsequent routine histological processing. Regions of interest (ROIs) of cellular areas were outlined in D-FFOCT slides in ImageJ software. For each ROI, we quantified both the area and the color parameters. The color parameter was defined as the average of red, green, and blue values across all pixels within each ROI.

Labels and human diagnosis at the slide level

FFPE blocks served as the unit for slide-level diagnosis, and one D-FFOCT image was acquired per block. Two experienced pathologists established the pathological gold standard diagnosis for each FFPE block. Immunohistochemical staining was performed to confirm the results in equivocal cases. This consensus diagnosis constituted the slide-level ground-truth label.

Two breast surgeons experienced and trained in D-FFOCT interpretation read and scored all D-FFOCT slides, which were then evaluated for diagnostic accuracy according to slide-level labels. The scoring system, as described in our previous article[15], used a 4-point scale (0: normal; 1: probably benign; 2: highly suspicious of malignancy; 3: probably malignant). Scores of 2 or 3 were considered positive for malignancy, whereas scores of 0 and 1 were considered negative for malignancy (**Figures S1 and S2**). A consensus was reached by discussion between the readers of D-FFOCT images with different scores. For false-positive cases, serial sections of FFPE blocks were prepared and stained with H&E. If malignancy was identified in serial sections, the slide-level labels were revised to positive for subsequent analyses.

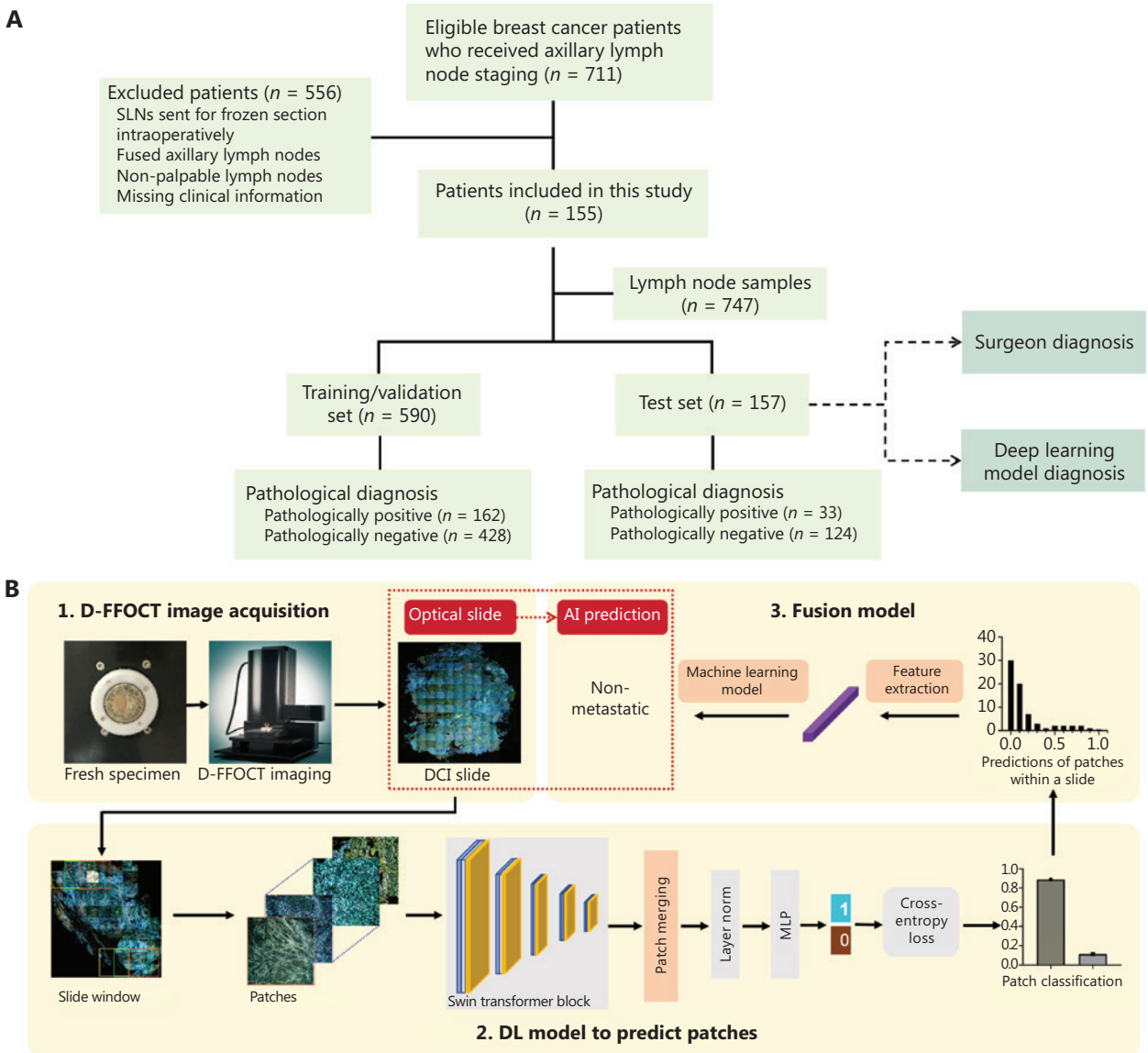


Figure 1 Flow of participants and deep learning model algorithm used throughout the study. (A) A total of 711 patients were eligible. Subsequently, 551 patients with matted or non-palpable lymph nodes or SLNs that required immediate FS and with missing clinical information were excluded, and 155 patients were ultimately enrolled. A total of 747 fresh lymph node specimens were obtained, and 747 D-FFOCT slides were acquired intraoperatively. (B) A two-step approach was used, wherein the pretrained Swin-T model first encoded the input D-FFOCT patches to a classification result, after which the combined features of a given D-FFOCT slide were classified by a machine learning fusion model. The goal of the overall pipeline was to output the D-FFOCT image at the slide level, as shown in the red dashed box.

Patch labeling and dataset

D-FFOCT slides were processed with a sliding window algorithm (1,150 × 1,150 pixels; 1,150-pixel stride) to generate small patches. The resulting patches were downsampled to 224 × 224 pixels to accommodate the deep learning model. Lymph node architecture and metastatic foci were segmented on D-FFOCT slides with QuPath[17] by pathologists with expertise in breast histology. Only patches containing ≥ 20% lymph node tissue were retained. Patch labels were determined according to the presence of metastatic pixels within the patches.

On the basis of the patient enrollment timeline, the patches were allocated to training/validation and independent test sets in a 4:1 ratio. All samples from Beijing Chaoyang Hospital

of Capital Medical University were assigned to the test set. A 5-fold cross-validation approach was used by dividing the training/validation set into 5 segments.

Deep learning framework and visualization

We based our model on a tiny version of the Swin Transformer (Swin-T) architecture, which demonstrated superior performance for D-FFOCT breast imaging in previous research[16]. The entire pipeline is shown in **Figure 1B**. Model parameters were initialized with ImageNet-1K pretrained weights, and fine-tuning was subsequently performed *via* backpropagation. Cross-entropy loss was used when the model weights were adjusted during training. The optimizer used Adam with a batch size of 64, a learning rate of e-6, and training of 30 epochs. The model was implemented in Python 3.7, which is based on the PyTorch deep learning library, and the neural networks were trained on a workstation with a GeForce GTX3090 GPU (Nvidia). During the prediction stage, the patches were passed through the trained network for classification. The average of the 5 output values was considered the final output for each patch during testing. The method used to align the patch prediction results with the slide levels was similar to that described in our previous report[16]. To enhance model interpretability, we applied gradient-weighted class activation mapping (Grad-CAM) to generate activation heatmaps highlighting important regions of the image for the classification task.

Statistical analysis

We hypothesized that the diagnostic ability of the deep learning model would be noninferior to that of surgeons in the interpretation of D-FFOCT images of lymph nodes. The noninferiority margin was based on the findings that artificial intelligence (AI)-supported diagnosis is considered noninferior when its accuracy is approximately 5%–10% lower than that of human experts[18,19]. Therefore, a 7% noninferiority margin was agreed upon at a consensus group meeting including breast surgeons and pathologists (Shuwei Zhang, Houpu Yang, and YYZ). This noninferiority trial was designed with the following parameters: expected accuracy of 94%, a noninferiority margin of 7%, an α (alpha level) of 0.05, and a statistical power of 0.9. On the basis of these parameters, we calculated a required sample size of 156 slides in the test set. The sensitivity, specificity, false-negative rate, false-positive rate, and accuracy were calculated at the slide level. Receiver operating characteristic curve analysis was performed to evaluate the performance of the surgeons and the AI model. DeLong's test was applied to evaluate the differences in diagnostic performance between surgeons and the model. Subgroup diagnostic accuracy comparisons were performed with χ^2 or Fisher's exact tests. All tested hypotheses were 2-sided, and a P value less than 0.05 was considered significant.

Results

Patients and the dataset

From May 2018 to October 2022, 155 patients with breast cancer were enrolled in the study. A total of 747 axillary lymph node D-FFOCT slides were generated and included in the analysis (**Table 1**). The imaging process was tissue preserving and did not include contrast or staining. All experimental procedures, except for the pathological diagnosis, including D-FFOCT imaging and categorization, were completed independently by the surgeons. In total, 590 slides (28,911 patches) were included in the training/validation set, and 157 slides (7,736 patches) were included in the test set. The test set contained 33 slides (1,823 patches) classified as malignant and 124 slides (5,913 patches) classified as benign. The distribution of malignant slides in both sets was balanced.

Near-histological imaging of lymph nodes by D-FFOCT

In this expanded cohort, D-FFOCT also demonstrated high concordance with H&E-based histopathology at both the tissue and cellular levels for the discrimination of nonmetastatic and metastatic lymph nodes. Representative imaging (**Figure 2A**) revealed cortical (outer capsule and lymphoid follicles) and medullary (medullary cords) structures. Metastatic nodes exhibited structural effacement, collagen tangles, and clustered yellow malignant cells. D-FFOCT clearly distinguished metastatic patterns across tumor subtypes. Mucinous carcinoma metastases exhibit characteristic mucin pools containing floating tumor cell nests, which are distinct from those of invasive ductal carcinoma. D-FFOCT enabled virtual histological visualization of micrometastases and reliably distinguished between malignant and normal cells in smaller metastatic lesions (**Figure 2B**).

Table 1 Characteristics of included slides and patches

Characteristics	Total no.	Training/ validation set	Test set
D-FFOCT slides included	747	590	157
With metastasis	195	162	33
Macrometastasis	180	150	30
Micrometastasis and ITC	15	12	3
Without metastasis	552	428	124
Patches	36,647	28,911	7,736
With metastasis	10,800	8,977	1,823
Without metastasis	25,847	19,934	5,913

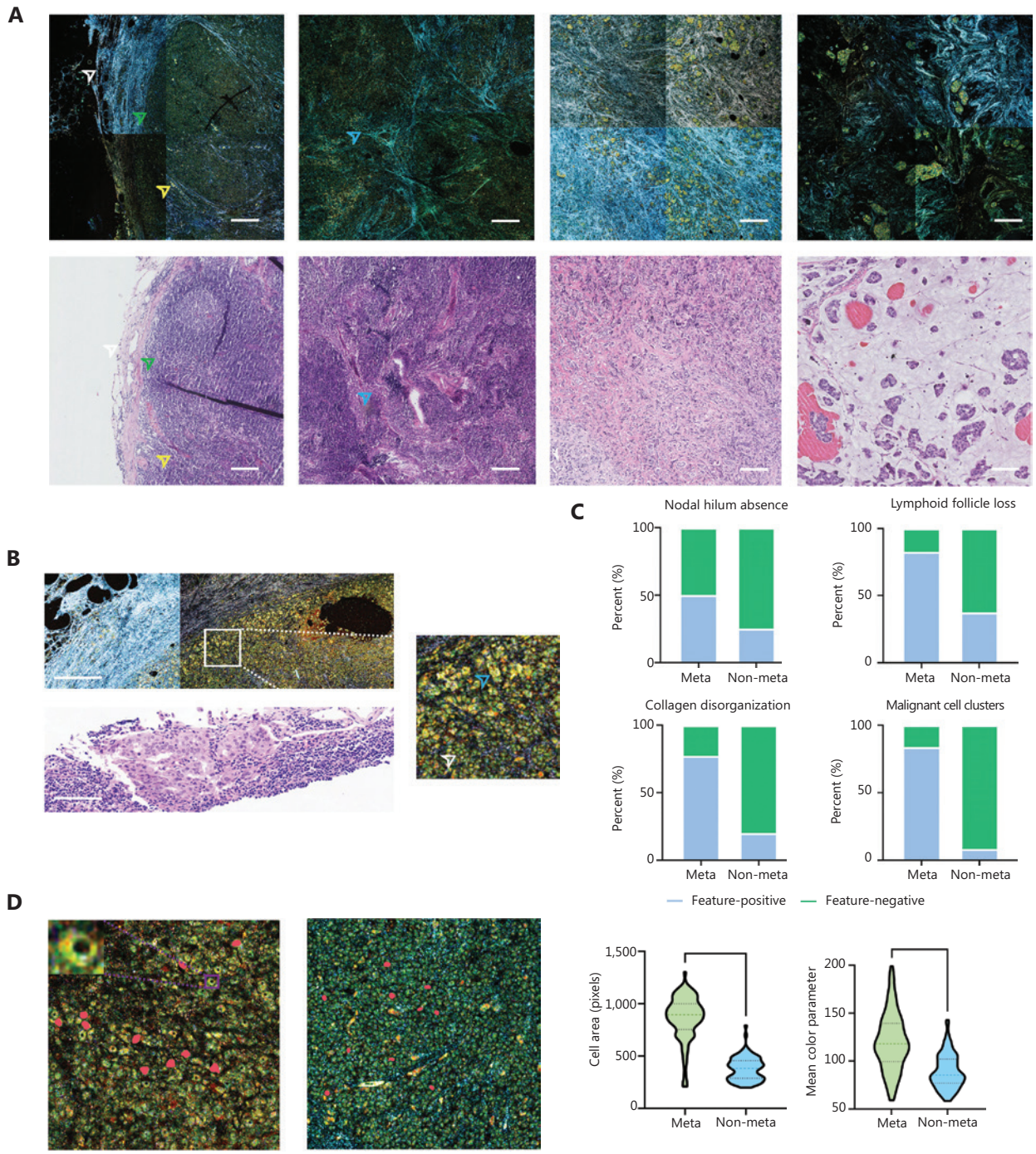


Figure 2 Near-histological imaging of the lymph nodes by D-FFOCT. (A) Representative D-FFOCT images and corresponding H&E histology of the lymph node cortex, medulla, invasive ductal carcinoma with metastasis, and mucinous carcinoma with metastasis. The white, green, and blue arrows indicate the outer capsule, lymphoid follicles, and medullary cords, respectively. Yellow arrows indicate trabeculae, which divide lymph nodes into compartments. (B) Visualization of micrometastasis by D-FFOCT and H&E. In the magnified view, larger tumor cells (blue arrow) contrast with small normal lymphocytes (white arrow). (C) Comparison of metastatic *vs.* non-metastatic lymph node features on D-FFOCT. (D) Quantitative comparison of cellular size and chromatic features between lymphocytes and tumor cells. Scale bar = 250 μm.

Trained surgeons identified 4 key tissue landscapes indicative of metastasis: nodal hilum absence, lymphoid follicle loss, collagen disorganization, and malignant cell clusters (**Figure 2C**). Although all features demonstrated diagnostic utility, the honeycomb patterns of malignant cells emerged as the most robust diagnostic marker (83.7% *vs.* 8.3% in

metastatic *vs.* non-metastatic nodes, $P < 0.05$). For the quantitative analysis, 10 ROIs were randomly sampled from both groups, with 10 cells analyzed per ROI (**Figure 2D**). Compared with metastatic tumor cells, normal lymphocytes exhibited tighter spatial organization, smaller cell areas, and distinct color signatures. Central black circular structures, putatively representing nucleoli, were occasionally observed in malignant cells.

Diagnostic performance of the human experts

During the initial diagnosis, 20 D-FFOCT slides were classified as false positives. Thereafter, serial sections of the FFPE blocks from these samples, with an average of 200 sections per block, revealed carcinoma in 4 cases. Therefore, the labels of these 4 specimens were modified to be listed as positive in the final diagnostic report.

Of the surgical diagnoses for the 747 slides from the entire dataset, the overall sensitivity and specificity were 82.56% [161/195, 95% confidence interval (CI): 76.34%–87.47%] and 97.10% (536/552, 95% CI: 95.23%–98.28%), respectively. Sixteen samples had a false-positive diagnosis: 15 classified as category 2 and 1 classified as category 3. The overall accuracy of the diagnoses by trained surgeons was 93.31% (697/747, 95% CI: 91.28%–94.89%), and the area under the receiver operating characteristic curve (AUC) was 0.918 (95% CI: 0.889–0.951) (**Figure 3A**). For the slides in the test set ($n = 157$), the AUC for the human readers was 0.919 (95% CI: 0.894–0.938), and the sensitivity and specificity were 78.79% (95% CI: 60.60%–90.37%) and 97.58% (95% CI: 92.56%–99.37%), respectively.

Performance of the deep learning algorithm

The training and validation loss curves, along with the patch-level accuracy, are shown in **Figure S3**. We used t-SNE to show class separation based on the feature representation learned by the final model classifier, as shown in **Figure 3B**. Clustering revealed strong class separation between 2 classes.

At the slide level, the difference in accuracy between the AI and human readers remained within the prespecified non-inferiority margin. The model correctly classified 91.08% (143/157, 95% CI: 85.59%–94.61%) of the lymph nodes in the test set, with a sensitivity and specificity of 87.88% (95% CI: 70.86%–96.04%) and 91.94% (95% CI: 85.30%–95.85%, **Figure 3C**), respectively, and an AUC of 0.899 (95% CI: 0.868–0.928). The slide-level receiver operating characteristic (ROC) curves of both the deep learning model and the human experts for the test set are shown in **Figure 3D**. DeLong's test revealed no statistically significant difference ($P = 0.627$). Notably, 11 slides from the external queue (second center) in the test set were correctly diagnosed.

To investigate the model's decision-making process, we implemented Grad-CAM to visualize malignancy-predictive regions on D-FFOCT images. The saliency maps consistently highlighted malignant cell clusters within lymphocyte backgrounds (**Figure 3E**). Notably, the model reliably identified tumor regions regardless of the brightness of malignant cells, thus suggesting sensitivity to structural alterations in both tumor and peritumoral areas. This finding demonstrates the robust feature extraction capability of the model.

Optical virtual histology workflow and subgroup diagnostic results

False-positive cases warrant particular concern in clinical scenarios, because they can lead to unnecessary tissue excisions. Although the diagnostic performance of the deep learning model was comparable to that of humans, trained surgeons demonstrated superior specificity (97.58% *vs.* 91.94%). We therefore established an integrated D-FFOCT workflow requiring review and final diagnostic determination by surgeons for all model-positive cases. In this real-world cohort, this approach necessitated human intervention in 24.84% of cases and delivered enhanced diagnostic results (**Figure 4A**). Implementation of this hybrid diagnostic pipeline yielded a 6.5% improvement in specificity, to 98.39% (95% CI: 93.71%–99.72%), and an overall accuracy of 93.63% (95% CI: 88.67%–96.50%).

This diagnostic paradigm demonstrated improved lymph node assessment across nearly all subgroups, including various tumor types, molecular subtypes, stages, and cases that received postneoadjuvant chemotherapy. However, the more stringent positivity criteria decreased the sensitivity and accuracy for the metastatic lymph node subgroups. Notably, in the micrometastasis and isolated tumor cell (ITC) subgroups ($n = 3$), the model did not identify any metastasis and consequently did not trigger the human confirmation workflow. Statistical testing indicated a significant difference in diagnostic performance between these subgroups ($P < 0.05$). Trained surgeons correctly diagnosed 40% (6/15) of cases in these challenging subgroups.

Discussion

D-FFOCT effectively differentiated between metastatic and nonmetastatic lymph nodes in this large-scale cohort. Moreover, this approach enabled virtual pathological imaging across different tumor types and metastatic burdens, thus demonstrating its broad clinical utility. The deep learning model exhibited robust performance in the independent test set, with a diagnostic AUC of 0.899. A combined human-AI interpretation strategy achieved a specificity of 98.39% and an overall accuracy of 93.63% while decreasing the manual diagnostic workload by 75%. This technology therefore has potential in streamlined lymph node evaluation during cancer surgery, by decreasing clinician workload and the reliance on routine pathology procedures (**Figure 4B**).

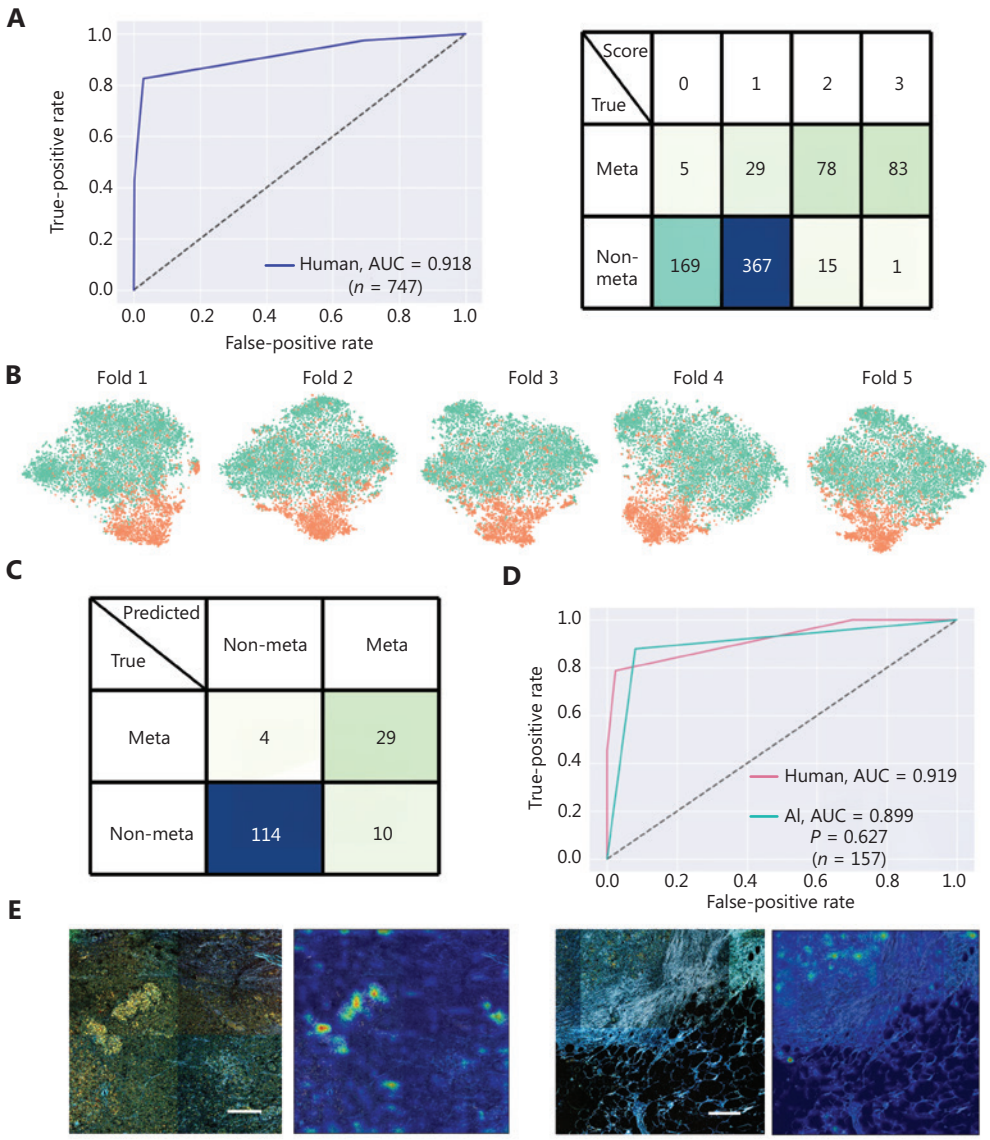


Figure 3 Performance of human experts and the deep learning model. (A) Diagnoses by human experts, including the ROC curve and confusion matrix, among all slides. (B) t-SNE plots of feature representations for the patches in the test set. (C) Confusion matrix of the deep learning model in the test set at the slide level. (D) Comparison of ROC curves between the model and human expert predictions in the test set. (E) D-FFOCT images with Grad-CAM heatmaps showing 2 malignant cell patterns: bright morphology (left) and atypical presentation (right). Scale bar = 250 μm.

Frozen sectioning is the most widely used method for rapid histopathological assessment of lymph nodes, and touch imprint cytology is occasionally used[5,20,21]. However, these methods remain suboptimal because of their time, tissue, and resource consumption, which are compounded by interobserver interpretive variability among pathologists[22-25]. Molecular approaches such as quantitative real-time reverse polymerase chain reaction and one-step nucleic acid amplification detect metastasis *via* tumor-specific mRNA in homogenized lymph nodes[26,27]. However, these methods face prohibitive cost barriers and cause irreversible tissue destruction precluding further analyses[28]. Whereas existing methods face persistent constraints, D-FFOCT has recently emerged as a promising method because of its unique advantages: label-free cellular imaging, preservation of specimen integrity, and histopathology-comparable accuracy (**Table 2**)[15,16,29,30]. In this study, we explored its application in nodal diagnosis in a large-scale sample.

Although trials such as Z0011 and SENOMAC have increased the omission of axillary dissection for selected SLN-positive patients with breast cancer[5,31-33], they underscore the need to accurately identify intervention candidates and to avoid missing clinically relevant metastases. Building on the ability

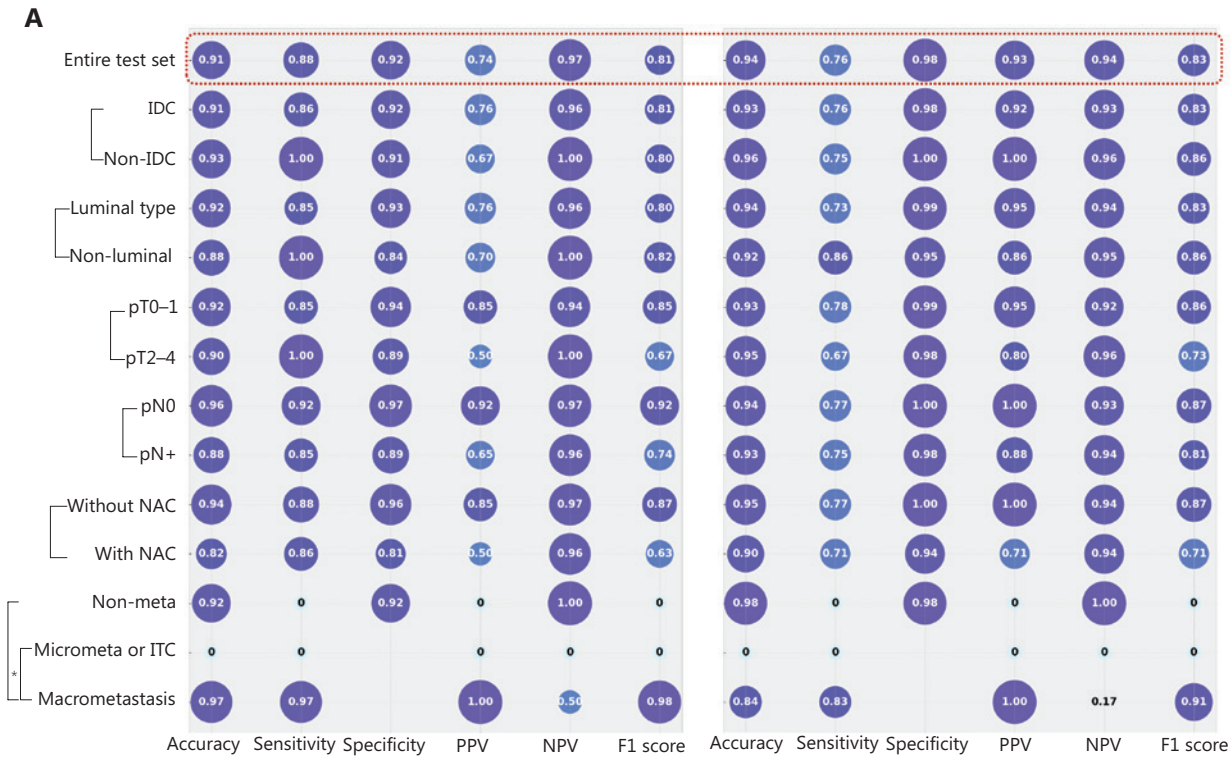


Figure 4 Continued

of D-FFOCT to be applied in primary tumors[15,34], this study extended its application to precise nodal evaluation. Our findings demonstrated that metastatic lymph nodes across tumor types consistently generated D-FFOCT images comparable to those of H&E-stained pathology images, thus confirming the applicability of this method in multiple scenarios. Simon et al. have performed complementary validation for biopsies[35] combined with easily learned image features and system portability. Their findings support the utility of D-FFOCT in diverse diagnostic contexts (including biopsies, primary tumor resection, margin assessment, and lymph node evaluation), by enabling optimized tumor diagnosis workflows across solid tumors.

Herein, reexamination of false-positive cases of D-FFOCT with deeper serial sections revealed occult metastases in 20% (4/20) of the FFPE blocks; these findings therefore confirmed the diagnostic validity of D-FFOCT image interpretation. This result might be attributable to limitations in the tissue sampling procedures of pathology. Previous studies have also confirmed that step sectioning can lead to the detection of metastases in lymph nodes that are initially determined to be negative[36]. The imaging principles of D-FFOCT endow cells in different metabolic states with different saturation and color, and demonstrate high fidelity in capturing tumor microarchitecture. Both cellular morphology and spatial arrangement were shown to be diagnostically effective (quantification in **Figure 2**). Bright yellow malignant cells demonstrated conserved optical features in samples with an oncological diagnosis that strongly correlated with H&E histology. This visualization therefore provides a reliable diagnostic tool for clinicians.

Although AI-assisted analysis of frozen sections may address labor-intensive aspects, this relatively nascent field has achieved suboptimal outcomes[37]. One study using automated deep learning for SLN diagnosis on frozen sections has reported a maximum accuracy of 77% in an independent test set of 100 slides[38]. In contrast, our D-FFOCT-based model achieved an accuracy of 91.08% in the test set. Swin-T was selected as the backbone, because its hierarchical design effectively captures multiscale features with relatively low computational cost[39-41]. The features learned from breast D-FFOCT images also appear to be conserved characteristics in samples with an oncological diagnosis, including the detection of lymph node metastases[16]. Critically, D-FFOCT also provides slide-free and tissue-preserving imaging for bedside clinical use. Nevertheless, image stitching artifacts remain a factor influencing AI model performance (**Figure S4**). Ongoing advances in D-FFOCT imaging, such as singular value decomposition-based filtering[42] and active phase modulation[43], are expected to further improve color uniformity.

The success of the CAMELYON challenges has demonstrated robust deep learning capabilities for automated

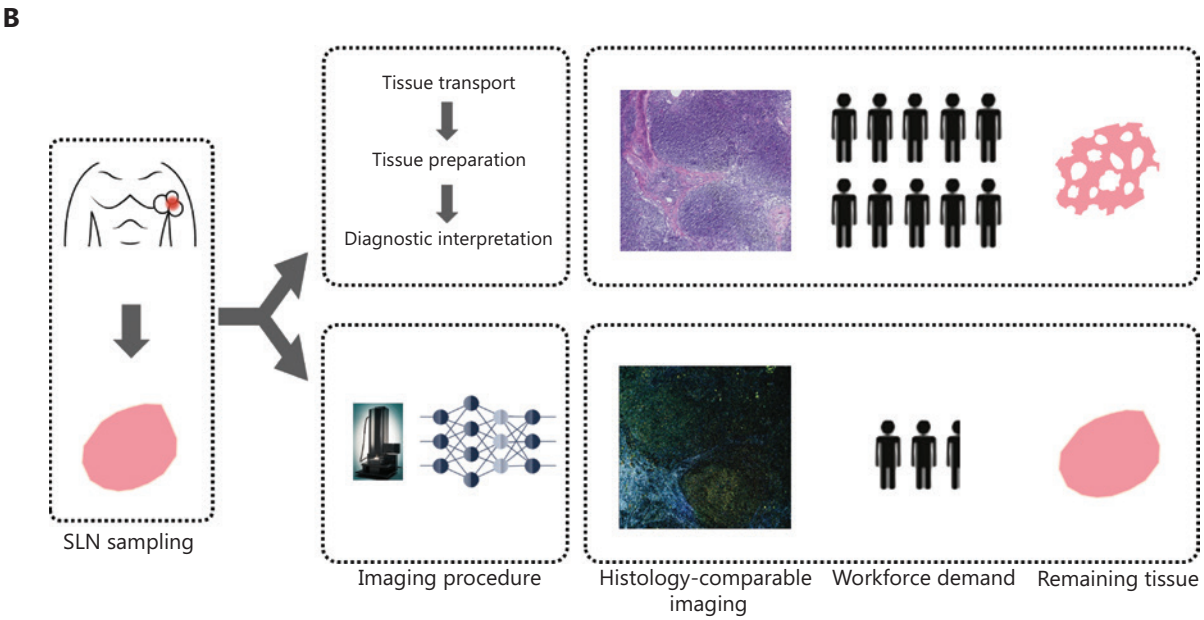


Figure 4 Optical virtual histology workflow and subgroup diagnostic results. (A) Bubble plot visualization of various performance indicators of deep learning models (left) and the human-AI collaborative paradigm (right) in the test set and different subgroups. Statistical significance in subgroup comparisons was evaluated for the human-AI collaborative workflow. A significant difference was observed only between the micrometastases and isolated tumor cells (ITC) subgroups ($P < 0.05$), as indicated. (B) Streamlined diagnosis workflows for lymph node assessment based on D-FFOCT plus deep learning. Our pipeline substantially decreased labor demands, eliminated tissue preparation requirements, and maintained specimen integrity for postoperative examinations. PPV, positive predictive value; NPV, negative predictive value.

Table 2 Comparison of nodal assessment methods

	D-FFOCT + AI	Frozen section	Molecular methods
Resolution	Approximately micrometer	Approximately micrometer	–
Time	~20 min	~30 min	~30 min
Tissue preservation	Fully tissue preserving	Tissue consuming	Tissue consuming
Labor consumption	Low	High	Low
Space and equipment costs	Moderate	High	Very high

diagnosis in H&E-stained lymph nodes[44,45]. Our results showed that D-FFOCT enables virtual pathological and high-contrast imaging between normal and malignant cells, and facilitates rapid learning. Recently, our collaborative work has established structural correspondence between optical and pathological imaging and developed generative models for D-FFOCT to synthesize H&E histology[46]. Because native images maximize diagnostic information retention in clinical settings, these findings motivated our direct diagnostic application of optical SLN images. Our results confirmed the features of D-FFOCT images and supported efficient model learning. Independent validation across diverse tumor types and stages consistently demonstrated high accuracy, thus aiding in establishing a viable automated diagnostic solution for SLN assessment.

This study revealed a potential synergy between human experts and AI models. Given the higher specificity of human-based diagnosis, a combined AI-expert approach might have broad appeal, because it increased the specificity to 98.39% and would decrease unnecessary aggressive surgeries. Our findings suggested that better outcomes might result from cooperation between the model and a surgeon than from either the AI model or surgeon alone. Because of the substantial labor savings provided by use of AI, future cloud-based platforms could help increase the expertise of D-FFOCT interpretation in resource-limited settings to a level similar to that of specialists. Two false-positive cases identified by our collaborative workflow occurred among patients after neoadjuvant chemotherapy. The D-FFOCT images revealed relatively large and regularly rearranged cells that misled the experts and the model (**Figure S5**). We hypothesized that this phenomenon might represent chemotherapy-induced collagenization of tumor cells and histiocyte aggregation. However, the limited inclusion of such samples resulted in insufficient learning of these features.

This study has several limitations. First, potential biases existed in the study design, including the temporal split of datasets, as well as the sampling discrepancy between our methods and real-world SLNB practice. Second, because only D-FFOCT false-positive cases underwent serial sectioning, sensitivity might potentially have been underestimated. Third, this work was an exploratory investigation of a novel technique and had a limited sample size. The small cohort from the second center warrants further validation, and the diagnostic performance for micrometastases and isolated tumor cells remains limited, in agreement with our previous findings[16]. Given the small sample size in this subgroup (n = 3), these results are exploratory and should be interpreted with caution. We plan to integrate additional samples of small tumor foci to reevaluate model performance and further improve diagnostic accuracy through model optimization. Despite these limitations, this study demonstrated the feasibility of D-FFOCT- and AI-based lymph node diagnosis. Future improvements in image quality, system stability, and sample size expansion may enable more robust conclusions to be drawn. Notably, although the commercial D-FFOCT device lacks quantitative analysis functions, our self-built system introduced active modulation to achieve decoupled dynamic feature extraction, providing quantitative dynamic information unavailable in commercial devices[43]. Ongoing developments in three-dimensional and *in vivo* D-FFOCT imaging may further expand the clinical potential of this technology.

Conclusions

In conclusion, D-FFOCT combined with a deep learning algorithm provided a nonconsumptive approach to evaluating lymph node metastasis in patients with breast cancer. This approach decreases labor demands and provides accurate assistance in surgical decisions. In the future, we plan to extend the application of D-FFOCT to patients with other tumor types.

Acknowledgements

We gratefully acknowledge Professor Xiaoxia Peng at Beijing Children's Hospital, Capital Medical University, for providing invaluable statistical consultation and expert guidance during the revision of this manuscript.

Grant support

This work was supported by grants from the National Key Research and Development Program of China (Grant No. 2024YFC3405303), Beijing Natural Science Foundation (Grant No. 7242281 and 7244427), and Research and Development Fund of Peking University People's Hospital (Grant No. RDZH2024-03 and RDEB2025-25).

Conflict of interest statement

No potential conflicts of interest are disclosed.

Author contributions

Conceived and designed the study: Shu Wang.
Performed the experiments and collected the data: Shuwei Zhang, Houpu Yang, Yiyin Zhang, and Jin Zhao.
Performed pathological data curation and annotation of malignant regions in D-FFOCT images: Xiaoxian Li and Yuanyuan Zhang.
Performed statistical analysis and draft writing: Shuwei Zhang.
Reviewed and edited the manuscript: Houpu Yang, Yiyin Zhang, and Xiaoxian Li.
Provided technical and analytical support: Ping Xue.
Provided project resources: Hongchuan Jiang and Hua Kang.
Provided statistical consultation and advice: Wenhui Ren.

Data availability statement

Data used in this manuscript are available on request for non-commercial and academic purposes from corresponding author Shu Wang (shuwang@pkuph.edu.cn).

References

1. Park KU, Somerfield MR, Anne N, Brackstone M, Conlin AK, Couto HL, et al. Sentinel lymph node biopsy in early-stage breast cancer: ASCO guideline update. J Clin Oncol. 2025; 43: 1720-41.
2. Zhang X, Bao B, Wang S, Yi M, Jiang L, Fang X. Sentinel lymph node biopsy in early stage cervical cancer: a meta-analysis. Cancer Med. 2021; 10: 2590-600.
3. Sullivan SA, Rossi EC. Sentinel lymph node biopsy in endometrial cancer: a new standard of care? Curr Treat Options Oncol. 2017; 18: 62.
4. Garau LM, Rubello D, Ferretti A, Boni G, Volterrani D, Manca G. Sentinel lymph node biopsy in small papillary thyroid cancer. A review on novel surgical techniques. Endocrine. 2018; 62: 340-50.
5. Lombardi A, Nigri G, Maggi S, Stanzani G, Vitale V, Vecchione A, et al. Role of frozen section in sentinel lymph node biopsy for breast cancer in the era of the ACOSOG Z0011 and IBCSG 23-10 trials. Surgeon. 2018; 16: 232-36.
6. Li F, Yang L, Zhao Y, Yuan L, Wang S, Mao Y. Intraoperative frozen section for identifying the invasion status of lung adenocarcinoma: a systematic review and meta-analysis. Int J Surg. 2019; 72: 175-84.
7. Winter M, Gibson R, Ruszkiewicz A, Thompson SK, Thierry B. Beyond conventional pathology: towards preoperative and intraoperative lymph node staging. Int J Cancer. 2015; 136: 743-51.

8. Huang L, Sun H, Sun L, Shi K, Chen Y, Ren X, et al. Rapid, label-free histopathological diagnosis of liver cancer based on raman spectroscopy and deep learning. Nat Commun. 2023; 14: 48.
9. Pradipta AR, Tanei T, Morimoto K, Shimazu K, Noguchi S, Tanaka K. Emerging technologies for real-time intraoperative margin assessment in future breast-conserving surgery. Adv Sci (Weinh). 2020; 7: 1901519.
10. Park J, Bai B, Ryu D, Liu T, Lee C, Luo Y, et al. Artificial intelligence-enabled quantitative phase imaging methods for life sciences. Nat Methods. 2023; 20: 1645-60.
11. Apelian C, Harms F, Thouvenin O, Boccara AC. Dynamic full field optical coherence tomography: subcellular metabolic contrast revealed in tissues by interferometric signals temporal analysis. Biomed Opt Express. 2016; 7: 1511-24.
12. Ren C, Hao S, Wang F, Matt A, Amaral MM, Yang D, et al. Dynamic contrast optical coherence tomography (DyC-OCT) for label-free live cell imaging. Commun Biol. 2024; 7: 278.
13. Yang L, Yu X, Fuller AM, Troester MA, Oldenburg AL. Characterizing optical coherence tomography speckle fluctuation spectra of mammary organoids during suppression of intracellular motility. Quant Imaging Med Surg. 2020; 10: 76-85.
14. Scholler J, Groux K, Goureau O, Sahel JA, Fink M, Reichman S, et al. Dynamic full-field optical coherence tomography: 3D live-imaging of retinal organoids. Light Sci Appl. 2020; 9: 140.
15. Yang H, Zhang S, Liu P, Cheng L, Tong F, Liu H, et al. Use of high-resolution full-field optical coherence tomography and dynamic cell imaging for rapid intraoperative diagnosis during breast cancer surgery. Cancer. 2020; 126(Suppl 16): 3847-56.
16. Zhang S, Yang B, Yang H, Zhao J, Zhang Y, Gao Y, et al. Potential rapid intraoperative cancer diagnosis using dynamic full-field optical coherence tomography and deep learning: a prospective cohort study in breast cancer patients. Sci Bull (Beijing). 2024; 69: 1748-56.
17. Bankhead P, Loughrey MB, Fernández JA, Dombrowski Y, McArt DG, Dunne PD, et al. QuPath: open source software for digital pathology image analysis. Sci Rep. 2017; 7: 16878.
18. Eisemann N, Bunk S, Mukama T, Baltus H, Elsner SA, Gomille T, et al. Nationwide real-world implementation of ai for cancer detection in population-based mammography screening. Nat Med. 2025; 31: 917-24.
19. Hollon TC, Pandian B, Adapa AR, Urias E, Save AV, Khalsa SSS, et al. Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks. Nat Med. 2020; 26: 52-8.
20. Naritoku WY, Furlong MA, Knollman-Ritschel B, Kaul KL. Enhancing the pipeline of pathologists in the United States. Acad Pathol. 2021; 8: 23742895211041725.
21. Peng L, Zheng X, Xue Y, Huang C, Su X, Yu S. Central lymph nodes in frozen sections can effectively guide extended lymph node resection for papillary thyroid carcinoma. Ann Med. 2023; 55: 2286337.
22. Riedl O, Fitzal F, Mader N, Dubsky P, Rudas M, Mittlboeck M, et al. Intraoperative frozen section analysis for breast-conserving therapy in 1016 patients with breast cancer. Eur J Surg Oncol. 2009; 35: 264-70.
23. Jaafar H. Intra-operative frozen section consultation: concepts, applications and limitations. Malays J Med Sci. 2006; 13: 4-12.
24. Layfield DM, Agrawal A, Roche H, Cutress RI. Intraoperative assessment of sentinel lymph nodes in breast cancer. Br J Surg. 2011; 98: 4-17.
25. Tew K, Irwig L, Matthews A, Crowe P, Macaskill P. Meta-analysis of sentinel node imprint cytology in breast cancer. Br J Surg. 2005; 92: 1068-80.
26. Hao X, Liu Y, Li X, Kang H, Qu X, He J, et al. An intra-operative RT-LAMP method allows rapid and reliable detection of sentinel lymph node metastasis in breast cancer patients. Virchows Arch. 2015; 466: 169-76.
27. Shi F, Liang Z, Zhang Q, Wang C, Liu X. The performance of one-step nucleic acid amplification assay for intraoperative detection of sentinel lymph node macrometastasis in breast cancer: an updated meta-analysis. Breast. 2018; 39: 39-45.
28. van der Noordaa MEM, Vrancken Peeters MTFD, Rutgers EJT. The intraoperative assessment of sentinel nodes – standards and controversies. Breast. 2017; 34(Suppl 1): S64-9.
29. Park S, Nguyen T, Benoit E, Sackett DL, Garmendia-Cedillos M, Pursley R, et al. Quantitative evaluation of the dynamic activity of HeLa cells in different viability states using dynamic full-field optical coherence microscopy. Biomed Opt Express. 2021; 12: 6431-41.
30. Groux K, Verschueren A, Nanteau C, Clémençon M, Fink M, Sahel JA, et al. Dynamic full-field optical coherence tomography allows live imaging of retinal pigment epithelium stress model. Commun Biol. 2022; 5: 575.
31. Giuliano AE, Ballman KV, McCall L, Beitsch PD, Brennan MB, Kelemen PR, et al. Effect of axillary dissection vs no axillary dissection on 10-year overall survival among women with invasive breast cancer and sentinel node metastasis: the ACOSOG Z0011 (Alliance) randomized clinical trial. JAMA. 2017; 318: 918-26.
32. Harrison B. Update on sentinel node pathology in breast cancer. Semin Diagn Pathol. 2022; 39: 355-66.
33. de Boniface J, Filtenborg Tvedskov T, Rydén L, Szulkin R, Reimer T, Kühn T, et al. Omitting axillary dissection in breast cancer with sentinel-node metastases. N Engl J Med. 2024; 390: 1163-75.
34. Scholler J, Mandache D, Mathieu MC, Lakhdar AB, Darche M, Monfort T, et al. Automatic diagnosis and classification of breast surgical samples with dynamic full-field OCT and machine learning. J Med Imaging (Bellingham). 2023; 10: 034504.
35. Simon A, Badachi Y, Ropers J, Laurent I, Dong L, Da Maia E, et al. Value of high-resolution full-field optical coherence tomography and dynamic cell imaging for one-stop rapid diagnosis breast clinic. Cancer Med. 2023; 12: 19500-11.
36. Valla M, Gjertrud Mjønes P, Knopp S. Implementation of step sectioning in the examination of sentinel lymph nodes to improve the detection of micrometastases in breast cancer patients. APMIS. 2012; 120: 521-8.

37. Gorman BG, Lifson MA, Vidal NY. Artificial intelligence and frozen section histopathology: a systematic review. J Cutan Pathol. 2023; 50: 852-9.
38. Kim YG, Song IH, Lee H, Kim S, Yang DH, Kim N, et al. Challenge for diagnostic assessment of deep learning algorithm for metastases classification in sentinel lymph nodes on frozen tissue section digital slides in women with breast cancer. Cancer Res Treat. 2020; 52: 1103-11.
39. Wang Y, Yue X, Lou S, Feng P, Cui B, Liu Y. Gene Swin transformer: new deep learning method for colorectal cancer prognosis using transcriptomic data. Brief Bioinform. 2025; 26: bbaf275.
40. Cai H, Feng X, Yin R, Zhao Y, Guo L, Fan X, et al. Mist: multiple instance learning network based on Swin Transformer for whole slide image classification of colorectal adenomas. J Pathol. 2023; 259: 125-35.
41. Chen S, Zhao X, Liu Y, Xie H, Wu Z, Xie Y, et al. Dual-scale Swin Transformer via feature alignment and adversarial discrimination for retinopathy of prematurity diagnosis. IEEE J Biomed Health Inform. 2025; 29: 9013-26.
42. Scholler J. Motion artifact removal and signal enhancement to achieve in vivo dynamic full field OCT. Opt Express. 2019; 27: 19562-72.
43. Yin Z, Zhang S, He B, Yang H, Chen Z, Hu Z, et al. Optical biomarker of metabolism for breast tumor diagnosis: insights from subcellular dynamics. arXiv preprint. 2024; arXiv:2406.03798.
44. Ehteshami Bejnordi B, Veta M, Johannes van Diest P, van Ginneken B, Karssemeijer N, Litjens G, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. JAMA. 2017; 318: 2199-210.
45. Jarkman S, Karlberg M, Pocevičiūtė M, Bodén A, Bándi P, Litjens G, et al. Generalization of deep learning in digital pathology: experience in breast cancer metastasis detection. Cancers. 2022; 14: 5424.
46. Yin Z, He B, Ying Y, Zhang S, Yang P, Chen Z, et al. Fast and label-free 3D virtual H&E histology via active phase modulation-assisted dynamic full-field OCT. npj Imaging. 2025; 3: 12.

Cite this article as: Zhang S, Yang H, Zhang Y, Li X, Zhao J, Zhang Y, et al. Virtual histology imaging of lymph nodes *via* dynamic full-field optical coherence tomography and deep learning to differentiate metastasis. Cancer Biol Med. 2026; 23: 418-429. doi: 10.20892/j.issn.2095-3941.2025.0747