Deep learning-based prediction of nodal metastasis in lung cancer using endobronchial ultrasound

Tsukasa Ishiwata, MD, PhD,[a] Terunaga Inage, MD, PhD,[a] Masato Aragaki, MD, PhD,[a] Alexander Gregor, MD, PhD,[a] Zhenchian Chen, MD,[a] Nicholas Bernards, PhD,[a] Kamran Kafi, MD,[b] and Kazuhiro Yasufuku, MD, PhD[a]

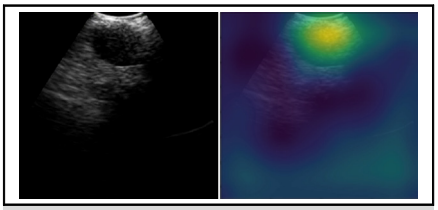
ABSTRACT

Objective: Endobronchial ultrasound-guided transbronchial needle aspiration is a vital tool for mediastinal and hilar lymph node staging in patients with lung cancer. Despite its high diagnostic performance and safety, it has a limited negative predictive value. Our objective was to evaluate the diagnostic performance of deep learning-based prediction of lung cancer lymph node metastases using convolutional neural networks developed from automatically extracted images of endobronchial ultrasound videos without supervision of the lymph node location.

Methods: Patient and lymph node data were collected from a single-center database. The diagnosis of metastasis was confirmed with endobronchial ultrasound-guided transbronchial needle aspiration and/or surgically resected specimens; the diagnosis of normal lymph node was confirmed with surgically resected specimens only. An annotation system facilitated automated image extraction from endobronchial ultrasound videos. Image frames were randomly selected and split into training and validation datasets on a per-patient basis. A deep learning model with convolutional neural networks, SqueezeNet, was used for image classification via transfer learning based on pretraining from ImageNet. Adaptive moment estimation and stochastic gradient descent were applied as optimizers.

Results: SqueezeNet, with adaptive moment estimation, achieved a sensitivity, specificity, accuracy, positive predictive value, and negative predictive value of 96.7% each after 300 epochs, whereas SqueezeNet with stochastic gradient descent achieved 91.1% each. However, SqueezeNet with stochastic gradient descent demonstrated more stable performance than with adaptive moment estimation.

Conclusions: Deep learning-based image classification using convolutional neural networks showed promising diagnostic accuracy for lung cancer nodal metastasis. Future clinical trials are warranted to validate the algorithm's efficacy in a prospective, large-cohort study. (JTCVS Techniques 2024;28:151-61)



Convolutional neural networks detect lymph node metastases with promising accuracy.

CENTRAL MESSAGE

Deep learning-based image classification using convolutional neural networks showed promising diagnostic accuracy for lung cancer nodal metastasis.

PERSPECTIVE

Endobronchial ultrasound-guided transbronchial needle aspiration is a vital tool for mediastinal/hilar lymph node staging in patients with lung cancer, but its negative predictive value is limited. Deep learning-based image classification using convolutional neural networks demonstrated promising diagnostic accuracy for nodal metastasis. Future clinical trials will be needed to validate this tool.

See Discussion on page 162.

From the [a]Division of Thoracic Surgery, Toronto General Hospital, University Health Network, Toronto, Ontario, Canada; and [b]Imagia Cybernetics, Montreal, Québec, Canada.

Funding provided through the William Coco Chair in Surgical Innovation for Lung Cancer.

Institutional Review Board approval: University Health Network, REB #19-5805, approved on October 17, 2019.

Informed consent statement: The study consent from participants was waived due to retrospective nature of the study.

Read at the 104th Annual Meeting of The American Association for Thoracic Surgery, Toronto, Ontario, Canada, April 27-30, 2024.

Received for publication April 29, 2024; revisions received Aug 15, 2024; accepted for publication Sept 11, 2024; available ahead of print Sept 19, 2024.

Address for reprints: Kazuhiro Yasufuku, MD, PhD, Division of Thoracic Surgery, Toronto General Hospital, University Health Network, 200 Elizabeth St, 9N-957, Toronto, Ontario M5G 2C4, Canada (E-mail: kazuhiro.yasufuku@uhn.ca).

2666-2507

https://doi.org/10.1016/j.xjtc.2024.09.008

Abbreviations and Acronyms

AI	= artificial intelligence
Adam	= adaptive moment estimation
EBUS-TBNA	= endobronchial ultrasound-guided transbronchial needle aspiration
Grad-CAM	= gradient-weighted class activation mapping
LN	= lymph node
SGD	= stochastic gradient descent
SVM	= support-vector machine

To view the AATS Annual Meeting Webcast, see the URL next to the webcast thumbnail.

Endobronchial ultrasound-guided transbronchial needle aspiration (EBUS-TBNA) has been widely performed in mediastinal and hilar lymph node (LN) staging in patients with lung cancer,[1,2] owing to its remarkable diagnostic efficacy and safety profile.[3-5] A meta-analysis demonstrated that EBUS-TBNA achieved a sensitivity of 90%, specificity of 99%, accuracy of 96%, positive predictive value of 99%, and negative predictive value of 93% in the staging of non–small cell lung cancer. However, the negative predictive value exhibits considerable variability across studies, ranging from 72% to 97%.[3] However, the negative predictive value exhibits considerable variability across studies, ranging from 72% to 97%.[6] It is important to note its limited negative predictive value, especially within cohorts exhibiting high pretest probability for LN metastasis.[6,7] The American College of Chest Physicians guidelines for lung cancer recommend that when there is a high pretest probability of N2 or N3 disease, a negative EBUS-TBNA result should prompt consideration of mediastinoscopy.[1] However, precise evaluation of pretest probability remains a challenge. In addressing this issue, some groups have sought to predict metastasis before sampling by analyzing EBUS images. The classification of LN characteristics within EBUS images has been explored as a potential strategy to distinguish between metastatic and benign LNs.[8,9] Nevertheless, it is important to acknowledge that this classification is inherently subjective and dependent on observer discretion. Previous research has demonstrated that grayscale texture analysis of EBUS images can effectively differentiate between malignant and benign LNs.[10,11] However, this approach has exclusively been conducted as a retrospective analysis rather than a real-time prediction tool during EBUS procedures.

The application of artificial intelligence (AI) technologies to ultrasound image recognition began with breast ultrasound images during the 1990s.[12] Presently, computer-aided detection systems demonstrate noteworthy diagnostic performance for ultrasound-based breast cancer detection.[13] Similarly, AI-based computer-aided detection/diagnosis has been implemented in endoscopic ultrasound for pancreatic and gastrointestinal tract diseases.[14-16]

Our future objective is to establish a real-time prediction system for metastatic LNs during EBUS procedures by leveraging AI technology to improve pretest probability estimation of nodal metastases. As the first step, this study focused on predicting metastatic LNs using AI based on images automatically extracted from EBUS videos, without supervision of LN location. We evaluated the diagnostic performance of convolutional neural networks as a deep learning-based image classification method for predicting lung cancer metastasis in LNs.

METHODS

Subjects

We retrospectively collected patients and corresponding LN data using the Interventional Thoracic Surgery Database at Toronto General Hospital, under the Institutional Review Board of the University Health Network (REB #19-5805; approved October 17, 2019). The study population comprised patients who underwent EBUS-TBNA for lung cancer staging at the Toronto General Hospital Interventional Thoracic Surgery Suite between January and December 2016. We included LNs in which a pathological diagnosis of lung cancer metastasis was made from EBUS-TBNA specimens and/or surgically resected specimens, or a pathological diagnosis of normal LNs was made from surgically resected specimens. LNs with inconclusive pathology results from EBUS-TBNA samples, such as atypia not diagnostic of malignancy, were not included if surgery samples were unavailable. Benign LNs with abnormal findings such as granuloma, fibrosis, and inflammation were not considered normal LNs. Exclusion criteria were surgically resected LNs showing a mix of metastatic and normal LNs within the same station, and LNs from patients who underwent any prior chemotherapy or chest radiotherapy. From the remaining LNs, we extracted those for which EBUS videos were available.

Procedures

EBUS-TBNA was performed per standard protocol,[17] using a BF-UC180 F (Olympus Medical Systems) bronchoscope and an EU-ME2 (Olympus Medical Systems) ultrasound processor. EBUS videos were recorded prospectively in an external hard drive in the endoscopy suite as is our usual clinical practice.

Automated Image Extraction From Videos

All EBUS videos were downloaded in the MP4 format and encoded with AAC, H.264 codec at 720 × 480 pixels after removing protected health information. The typical sequence during EBUS procedures consists of the following steps: identifying LNs, measuring the size of LNs using the caliper function, and labeling the LN station. This sequence was performed for LNs requiring assessment, followed by sampling via EBUS-TBNA. Building on this standard sequence, we developed a new annotation tool to automatically extract frames of interest from videos (Figure E1). During the measurement of LN size, the display zone showed blue lines representing a plus (+) for the first line and a cross (×) for the second line, along with the diameter displayed in the caliper zone, which the annotation system marked as PLUS and CROSS. When the labeling menu opened to input the station name of the LN, the annotation system marks it as MENU. The

sequence of events (PLUS-CROSS-MENU) was considered timestamps indicating the appearance of a LN of interest. Subsequently, 60 frames (2 seconds at 30 frames per second) before the start of the annotation process were automatically collected for each LN. Machine learning used only the display zone (400 × 375 pixels, depicted in an orange square in Figure E1) from the collected EBUS images.

Training and Validation Split

Collected image frames were split into the training dataset and the validation dataset, at 6:1 ratio, on a per-patient basis. To balance the number of metastatic LNs and normal LNs at 1:1, some image frames were randomly selected and removed from the datasets.

Image Processing

Basic pixel value normalization was applied (squeeze values between [0, 1]) and the display zone was cropped out, reducing the resolution by 50% (ie, 400 × 375 was scaled down to 200 × 188 pixels).

Models

As a nondeep learning model, the support-vector machine (SVM)[18] was used as a first step because SVM has shown good performances on imaging feature classification, including ultrasound images.[19-21] Features were obtained from a ResNet 50.[22] We then used principal component analysis, keeping 99% of the variance of the extracted features, and trained a simple SVM classifier with linear kernel to classify the malignancy of each frame.

As a deep learning model, we employed SqueezeNet, a deep convolutional neural network for image classification.[23] Transfer learning was applied using a pretraining model from ImageNet.[24] Adaptive moment estimation (Adam)[25] and stochastic gradient descent (SGD)[26] were utilized as optimizers and their performances were compared. Training was conducted using a learning rate of 10^{-4}, a momentum of 0.99, a batch size of 32, and epochs of up to 300. Image data augmentation (flipping or rotating) was not conducted. Additionally, gradient-weighted class activation mapping (Grad-CAM) was generated to localize the part that contributed to diagnosis.[27]

Statistical Analysis

Data were summarized as counts and medians with interquartile ranges for nonnormally distributed data. Sensitivity, specificity, accuracy, positive predictive value, and negative predictive value were calculated using standard definitions.

RESULTS

A total of 547 patients underwent EBUS-TBNA between January and December 2016, and among them, 129 patients with a total of 305 LNs met inclusion criteria. EBUS videos were available for 74 patients with a total of 133 LNs. LN detection in the EBUS videos using our annotation system was successful in 122 LNs (76 malignant, 46 normal), but failed in 11 LNs due to image artifacts and atypical sequences of the caliper and labeling flow. To balance the number of malignant and normal LNs equally, 11 videos were removed and the remaining 90 LNs from 53 videos were split into the training dataset (43 patients; 2340 frames of malignant LNs, 2340 frames of normal LNs) and the validation dataset (10 patients; 360 frames of malignant LNs, 360 frames of normal LNs) (Figure 1). The characteristics of extracted patients and LNs are shown in Table 1.

The diagnostic performance of the nondeep learning model using SVM was 84.4% (95% CI, 80.7%-88.0%)/45.2% (39.9%-50.6%)/66.1% (62.7%-69.6%)/63.8%

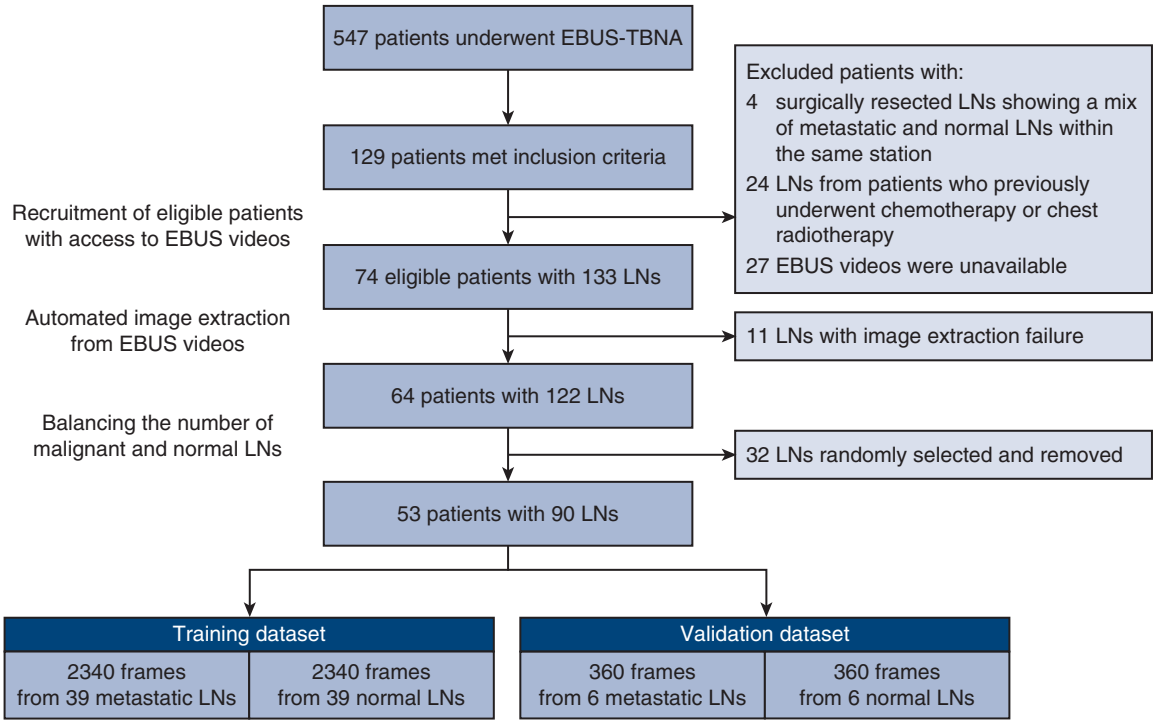


FIGURE 1. Consolidated Standards of Reporting Trials diagram for this study. *EBUS-TBNA*, Endobronchial ultrasound-transbronchial needle aspiration; *LN*, lymph node.

TABLE 1. The characteristics of study subjects

Characteristics	Overall		Training dataset		Validation dataset	
Patients	53		43		10	
Age (y)	69	(59-79)	71	(59-79)	64	(51-68)
Female sex	31	(58.5)	24	(55.8)	7	(70.0)
Lymph node diagnosis						
Metastatic lymph nodes						
Adenocarcinoma	19	(21.1)	15	(19.2)	4	(33.3)
Squamous cell carcinoma	6	(6.7)	5	(6.4)	1	(8.3)
Large cell lung cancer	2	(2.2)	2	(2.6)	0	(0)
NSCLC-NOS	8	(8.9)	8	(10.3)	0	(0)
Small cell lung cancer	4	(4.4)	3	(3.8)	1	(8.3)
Carcinoid	6	(6.7)	6	(7.7)	0	(0)
Normal lymph nodes	45	(50.0)	39	(50.0)	6	(50.0)
Lymph node size on EBUS images						
Metastatic lymph nodes						
Short axis (mm)	15.1	(10.5-20.9)				
Long axis (mm)	18.4	(13.5-23.8)				
Long-to-short axis ratio	1.15	(1.07-1.35)				
Normal lymph nodes						
Short axis (mm)	5.5	(4.0-8.8)				
Long axis (mm)	9.3	(6.5-13.1)				
Long-to-short-axis ratio	1.50	(1.23-1.83)				

	Training dataset				Validation dataset			
	Metastatic		Normal		Metastatic		Normal	
Lymph node image frames	2340		2340		360		360	
Station								
4R	780	(33.3)	780	(33.3)	60	(16.7)	60	(16.7)
4L	540	(23.1)	660	(28.2)	60	(16.7)	60	(16.7)
7	840	(35.9)	780	(33.3)	120	(33.3)	60	(16.7)
11Rs	120	(5.1)	60	(2.6)	60	(16.7)	60	(16.7)
11Ri	60	(2.6)	0	(0)	0	(0)	60	(16.7)
11L	0	(0)	60	(2.6)	60	(16.7)	60	(16.7)

Values are presented as n, n (%), or median (interquartile range). *NSCLC-NOS*, Non–small cell lung cancer not otherwise specified; *EBUS*, endobronchial ultrasound.

(59.6%-68.0%)/71.7% (65.6%-77.8%) in sensitivity/specificity/accuracy/positive predictive value/negative predictive value, respectively, whereas SqueezeNet with Adam after 300 epochs achieved 96.7% (95% CI, 94.8%-98.5%) for all performance metrics, and SqueezeNet with SGD after 300 epochs achieved 91.1% (95% CI, 88.2%-94.1%) for all performance metrics, respectively (Table 2). The area under the receiver operating characteristic curves based on LN size, LN long-to-short axis ratio, and each model are presented in Table E1. The learning process of SqueezeNet with Adam exhibited instability, whereas that of SqueezeNet with SGD demonstrated a reduction in prediction accuracy variance with an increasing number of epochs, resulting in a more stable convergence of the model (Figure E2).

Representative image frames for Grad-CAM are shown in Figure 2. In some images (Examples 1 and 2 in Figure 2), the areas identified by the trained model for prediction corresponded to the location of LNs within the image frames. However, there were also frames where the regions used for prediction did not align with the LN position (Examples 3 and 4 in Figure 2).

TABLE 2. Diagnostic performance comparison among prediction models

Prediction model	Sensitivity	Specificity	Accuracy	Positive predictive value	Negative predictive value
SVM (nondeep learning)	84.4 (80.7-88.0)	45.2 (39.9-50.6)	66.1 (62.7-69.6)	63.8 (59.6-68.0)	71.7 (65.6-77.8)
SqueezeNet with Adam	96.7 (94.8-98.5)	96.7 (94.8-98.5)	96.7 (94.8-98.5)	96.7 (94.8-98.5)	96.7 (94.8-98.5)
SqueezeNet with SGD	91.1 (88.2-94.1)	91.1 (88.2-94.1)	91.1 (88.2-94.1)	91.1 (88.2-94.1)	91.1 (88.2-94.1)

Values are presented as % (95% CI). *SVM*, Support-vector machine; *Adam*, adaptive moment estimation; *SGD*, stochastic gradient descent.

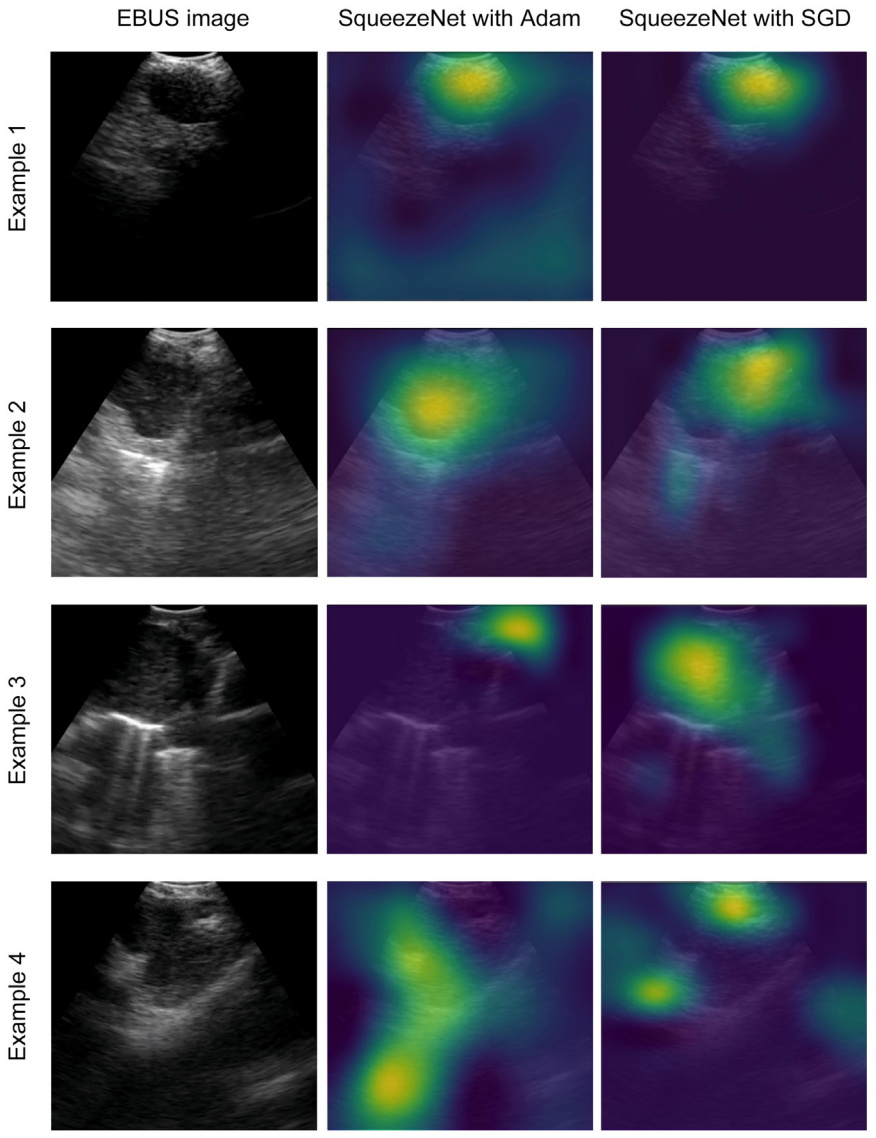


FIGURE 2. Gradient-weighted class activation mapping images in representative cases in the validation dataset. Example 1, station #7 lymph node with adenocarcinoma; example 2, station #11L with adenocarcinoma; example 3, station #11 Rs with adenocarcinoma; example 4, station #4R with squamous cell carcinoma. Examples 1 and 2 demonstrate accurate prediction in the lymph node location by SqueezeNet with adaptive moment estimation (*Adam*) and stochastic gradient descent (*SGD*) optimizers. Example 3 exhibits inadequate localization using SqueezeNet with Adam. Example 4 shows inadequate localization of the lymph node with SqueezeNet using both Adam and SGD. *EBUS*, Endobronchial ultrasound.

DISCUSSION

In our study, deep learning-based image classification employing convolutional neural networks on automatically extracted images from EBUS videos demonstrated promising diagnostic performance in identifying lung cancer metastases in mediastinal and hilar LNs (Figure 3). This pilot investigation is a proof of principle; nonetheless, beyond this, it may provide a fundamental dataset that could facilitate the development of a real-time LN predictive system during EBUS procedures.

The SqueezeNet model, when applied with a pretraining dataset, achieved a high diagnostic accuracy of 96.7% and 91.1% with Adam and SGD, respectively, in predicting LN metastasis within EBUS images in our study. Previous studies have explored various AI-based technologies for predicting LN diagnosis. In 2008, Tagaya and colleagues[28] reported the diagnostic performance of artificial neural networks in differentiating between LNs with sarcoidosis and those with lung cancer metastasis by using small regions of interest (32 × 32 pixels) from each EBUS image. Their study indicated superior performance of artificial neural networks compared to thoracic surgeons, with an accuracy of 91%. Similarly, Ozcelik and colleagues[29] utilized artificial neural networks on manually extracted regions of

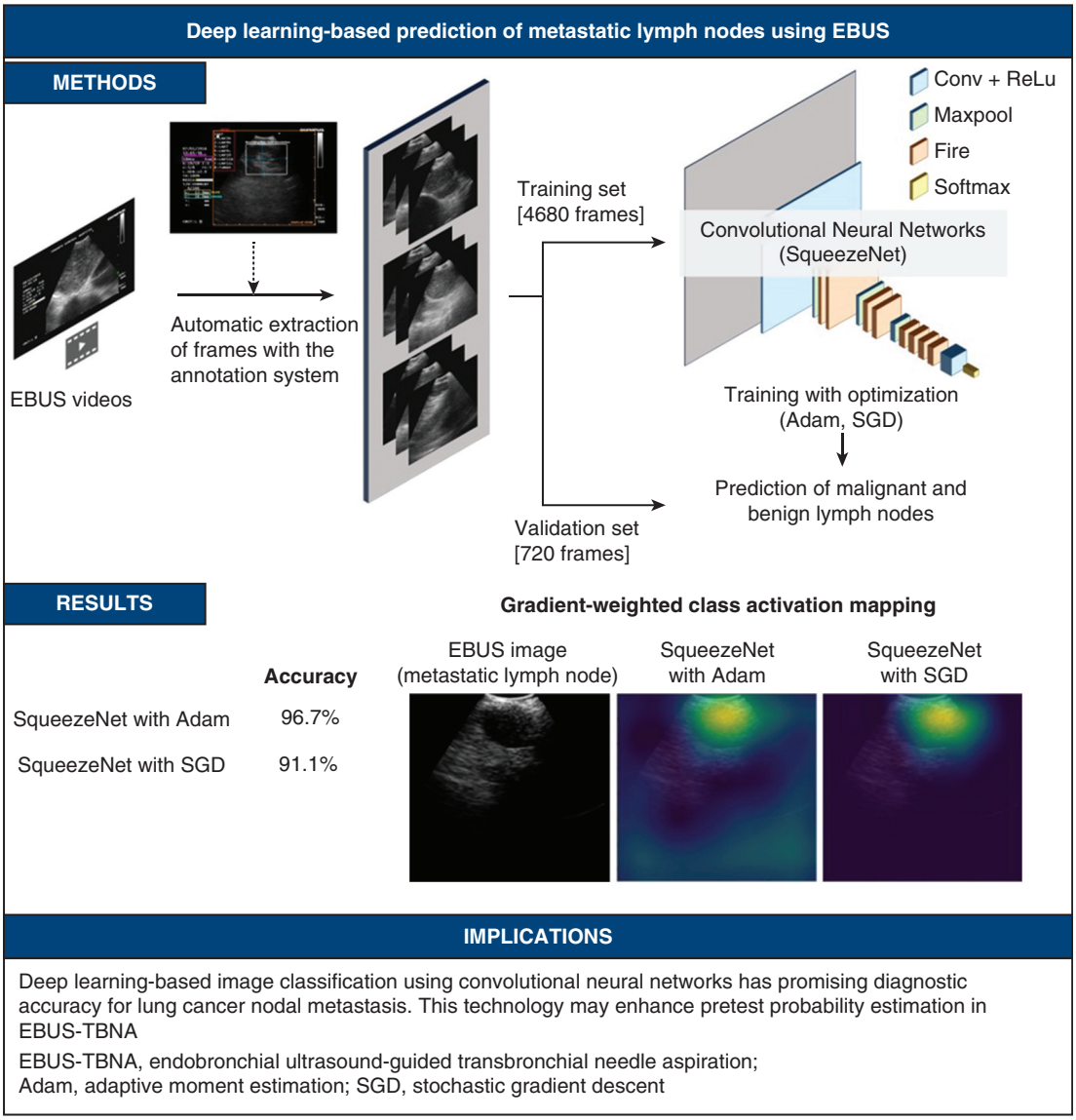


FIGURE 3. Overview of deep learning-based image classification utilizing convolutional neural networks for lung cancer nodal metastasis from endobronchial ultrasound (*EBUS*) images. We retrospectively collected data from patients with pathologically confirmed metastatic and normal lymph nodes. EBUS image frames were automatically extracted from recorded EBUS videos using a new annotation tool. A total of 4680 frames from 43 patients constituted the training dataset for a convolutional neural network model, SqueezeNet. Two different optimizers, adaptive moment estimation (*Adam*) and stochastic gradient descent (*SGD*), were employed during the learning process. The models were then tested using a validation dataset comprising 720 frames from 10 patients. SqueezeNet with Adam achieved a diagnostic accuracy of 96.7% for metastatic lymph nodes after 300 epochs, whereas SqueezeNet with SGD achieved 91.1%. The use of gradient-weighted class activation mapping revealed that the trained model accurately identified the location of lymph nodes within the image frames in some cases. This technology has the potential to improve the pretest probability of EBUS-guided transbronchial needle aspiration.

interest from EBUS images, achieving an 82% diagnostic accuracy for malignant and benign LNs. Yong and colleagues[30] evaluated convolutional neural networks using a large cohort, including 310 cases. Regions of interest on EBUS images were marked by bronchoscopists. The convolutional neural network achieved an accuracy of 75.8% in

predicting LN metastasis. However, these approaches relied on supervised selection of regions of interest, potentially limiting real-time applicability. In contrast, Churchill and colleagues[31] applied convolutional neural networks to entire EBUS images without selecting regions of interest, yielding a diagnostic accuracy of 72.9%. Ito and colleagues[32] also assessed diagnostic performance of convolutional neural networks using entire EBUS images, demonstrating a diagnostic accuracy of 87.9% for LN metastasis via a 5-fold cross-validation technique with the hold-out method. Li and colleagues[33] evaluated convolutional neural networks that combined features extracted from EBUS B-mode images, Doppler images, and elastographic images, demonstrating 88.6% accuracy. However, in clinical practice, this reliance on multiple modalities for all LNs could increase procedure time, certainly relative to systems that employ B-mode images alone.

Achieving similarly high real-world predictive performance for LN diagnosis as we have shown here would contribute to more accurately estimating the pretest probability before sampling of target LNs, resulting in improved lung cancer care. For instance, in patients with lung cancer where the pretest probability of metastatic LNs is high on EBUS images, the procedural approach may be modified to ensure greater tissue acquisition for genetic mutation screening[34,35] and programmed death-ligand 1 expression evaluation,[36] to appropriately determine the treatment strategy for advanced lung cancer and prevent delays in treatment initiation. Additionally, achieving high accuracy in predicting the diagnosis of individual LNs would lead to more efficient selection of nodes to puncture. Current guidelines for mediastinal staging by EBUS-TBNA require systematic nodal sampling from a minimum of 3 LN stations.[37] The selection of LNs for EBUS-TBNA sampling is often based on ease of access, such as larger nodes or those located closer to the bronchial wall. However, cases have been observed where positive and negative LNs coexist within the same LN station, potentially leading to false negatives in EBUS-TBNA.[38,39] In such cases, attempting selective sampling from LNs with higher pretest probability might reduce false negatives in EBUS-TBNA. Our ultimate goal of AI-assisted, EBUS-based prediction is to accurately estimate the prepuncture probability for malignancy to guide decision making following nondiagnostic EBUS-TBNA. In cases with inconclusive results from EBUS-TBNA, AI prediction might identify cases where repeat EBUS-TBNA or additional mediastinoscopy can be avoided. This could lead to a more efficient staging process in patients with lung cancer with potential benefits for cost-effectiveness. Future clinical trials should evaluate the incremental diagnostic benefit of additional AI support during EBUS-TBNA and assess the cost-effectiveness of introducing this technology.

Two different optimizers were employed with SqueezeNet in our study. Whereas SqueezeNet with Adam demonstrated high diagnostic accuracy, its unstable learning patterns indicated suboptimal model quality. Conversely, SqueezeNet with SGD demonstrated slightly lower diagnostic accuracy but maintained stable learning processes. Further evaluation of these developed algorithms is necessary in the context of prospective real-time assessment. Despite a recent study showing superior performance of fine-tuned machine learning methods like SVM over models including neural networks,[40] SVM in our study did not show sufficient predictive performance, achieving only 66.1% accuracy.

As demonstrated by Grad-CAM, in some LN images the focus of prediction accurately aligned with the LN itself rather than other mediastinal structures such as vessels. In this deep learning approach, only EBUS images and diagnosis results were provided; the supervision of LN location was not included. Nonetheless, the model still could accurately identify the positions of metastatic LNs that correspond with the actual LN locations. However, it should be noted, as shown in Figure 2, that there were image frames where the driver of the model's prediction was from an image region different from the true LN location.

This study has several limitations. First, this study was conducted at a single center with a limited number of participants and LNs, using a single EBUS processor. As a result, there was limited variation in LN stations, lung cancer subtypes, and benign LNs, potentially biasing the machine learning outcomes. Not all accessible EBUS stations were included in the training and validation datasets; only stations 4, 7, and 11 were used, whereas others (2, 3, 10, and 12) were not included. Structures other than LNs in these excluded stations, such as mediastinal tissues and vessels, may affect prediction accuracy. Furthermore, benign LNs with granulomas, fibrosis, and inflammation were not included; only normal LNs were selected in this study. The high positive/negative predictive values of our AI algorithms may therefore reflect the restrictions of our dataset. To address these limitations, training using a larger cohort encompassing a diverse range of malignant and benign LNs with testing of performance in a prospective clinical study is necessary. Secondly, micrometastases in LNs are unlikely to be detected with this AI approach, as the AI algorithm relies solely on EBUS B-mode images; micrometastases may be too small to generate architectural distortion detectable by EBUS. Although our cohort did not include any LNs with micrometastases, their presence should be carefully considered for future evaluation of AI in EBUS images. Thirdly, in our study, surgically resected LNs showing a mix of metastatic and normal nodes within the same station were excluded because it was challenging to correlate which LN corresponded with the EBUS video. This may have potentially overestimated the performance of AI prediction. Lastly, this study employed a

method of recognizing LNs of interest based on a specific labeling sequence on videos. This means that the frames containing LNs were selected under supervised conditions although the location of LNs was unsupervised.

Our objective is to establish a system capable of estimating metastatic LNs in real time from EBUS images. As the first step, we attempted to diagnose metastatic LNs using AI without supervising the location of LNs based on automatically extracted images from EBUS videos. In the next phase, our efforts will focus on developing a program using EBUS videos without any input of information from the bronchoscopist, creating a more flexible deep learning-based system to recognize frames containing LNs and predicting metastasis.

CONCLUSIONS

The deep learning-based image classification using convolutional neural networks in our pilot study demonstrated promising diagnostic performance for lung cancer nodal metastasis using automatically extracted images from EBUS videos. This achievement suggests the potential for deep-learning systems to provide real-time LN diagnosis predictions during EBUS-TBNA. Further evaluation of the developed algorithms in a prospective, large-cohort study is needed.

Webcast

You can watch a Webcast of this AATS meeting presentation by going to: https://www.aats.org/resources/deep-learning-based-prediction-7536.

Conflict of Interest Statement

Dr Yasufuku has industry-sponsored grants from Olympus Corp, Johnson and Johnson, and ODS Medical Inc; serves as a consultant for Olympus America Inc, Medtronic, Johnson and Johnson, and Astra Zeneca; collaborates on research with OKF Technology; and sits on the advisory board of Olympus America Inc, Medtronic, and Johnson and Johnson. All other authors reported no conflicts of interest.

The *Journal* policy requires editors and reviewers to disclose conflicts of interest and to decline handling or reviewing manuscripts for which they may have a conflict of interest. The editors and reviewers of this article have no conflicts of interest.

The authors are grateful to Philippe Lacaille, Emmanuel Jehanno, Francis Dutil, Tess Berthier, and Lisa Di Jorio for data preparation for analysis and program execution, and Judy McConnell for research coordination.

References

1. Silvestri GA, Gonzalez AV, Jantz MA, et al. Methods for staging non-small cell lung cancer: diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*. 2013;143(5 suppl):e211S-e250S.
2. De Leyn P, Dooms C, Kuzdzal J, et al. Revised ESTS guidelines for preoperative mediastinal lymph node staging for non-small-cell lung cancer. *Eur J Cardiothorac Surg*. 2014;45(5):787-798.
3. Dong X, Qiu X, Liu Q, Jia J. Endobronchial ultrasound-guided transbronchial needle aspiration in the mediastinal staging of non-small cell lung cancer: a meta-analysis. *Ann Thorac Surg*. 2013;96(4):1502-1507.
4. Yasufuku K, Pierre A, Darling G, et al. A prospective controlled trial of endobronchial ultrasound-guided transbronchial needle aspiration compared with mediastinoscopy for mediastinal lymph node staging of lung cancer. *J Thorac Cardiovasc Surg*. 2011;142(6):1393-1400.e1391.
5. Yu Lee-Mateus A, Garcia-Saucedo JC, Abia-Trujillo D, et al. Comparing diagnostic sensitivity of different needle sizes for lymph nodes suspected of lung cancer in endobronchial ultrasound transbronchial needle aspiration: systematic review and meta-analysis. *Clin Respir J*. 2021;15(12):1328-1336.
6. Whitson BA, Groth SS, Odell DD, et al. True negative predictive value of endobronchial ultrasound in lung cancer: are we being conservative enough? *Ann Thorac Surg*. 2013;95(5):1689-1694.
7. Czarnecka-Kujawa K, Yasufuku K. The role of endobronchial ultrasound versus mediastinoscopy for non-small cell lung cancer. *J Thorac Dis*. 2017;9(suppl 2):S83-S97.
8. Fujiwara T, Yasufuku K, Nakajima T, et al. The utility of sonographic features during endobronchial ultrasound-guided transbronchial needle aspiration for lymph node staging in patients with lung cancer: a standard endobronchial ultrasound image classification system. *Chest*. 2010;138(3):641-647.
9. Hylton DA, Turner S, Kidane B, et al. The Canada Lymph Node Score for prediction of malignancy in mediastinal lymph nodes during endobronchial ultrasound. *J Thorac Cardiovasc Surg*. 2020;159(6):2499-2507.e2493.
10. Nguyen P, Bashirzadeh F, Hundloe J, et al. Optical differentiation between malignant and benign lymphadenopathy by grey scale texture analysis of endobronchial ultrasound convex probe images. *Chest*. 2012;141(3):709-715.
11. Edey AJ, Pollentine A, Doody C, Medford AR. Differentiating benign from malignant mediastinal lymph nodes visible at EBUS using grey-scale textural analysis. *Respirology*. 2015;20(3):453-458.
12. Goldberg V, Manduca A, Ewert DL, Gisvold JJ, Greenleaf JF. Improvement in specificity of ultrasonography for diagnosis of breast tumors by means of artificial intelligence. *Med Phys*. 1992;19(6):1475-1481.
13. Wang X, Meng S. Diagnostic accuracy of S-Detect to breast cancer on ultrasonography: a meta-analysis (PRISMA). *Medicine (Baltimore)*. 2022;101(34):e30359.
14. Norton ID, Zheng Y, Wiersema MS, Greenleaf J, Clain JE, Dimagno EP. Neural network analysis of EUS images to differentiate between pancreatic malignancy and pancreatitis. *Gastrointest Endosc*. 2001;54(5):625-629.
15. Dumitrescu EA, Ungureanu BS, Cazacu IM, et al. Diagnostic value of artificial intelligence-assisted endoscopic ultrasound for pancreatic cancer: a systematic review and meta-analysis. *Diagnostics (Basel)*. 2022;12(2).
16. Liu XY, Song W, Mao T, Zhang Q, Zhang C, Li XY. Application of artificial intelligence in the diagnosis of subepithelial lesions using endoscopic ultrasonography: a systematic review and meta-analysis. *Front Oncol*. 2022;12:915481.
17. Nakajima T, Yasufuku K. The techniques of endobronchial ultrasound-guided transbronchial needle aspiration. *Innovations (Phila)*. 2011;6(1):57-64.
18. Cortes C, Vapnik V. Support-Vector Networks. *Mach Learn*. 1995;20(3):273-297.
19. Virmani J, Kumar V, Kalra N, Khandelwal N. SVM-based characterization of liver ultrasound images using wavelet packet texture descriptors. *J Digit Imaging*. 2013;26(3):530-543.
20. Baek J, Basavarajappa L, Hoyt K, Parker KJ. Disease-specific imaging utilizing support vector machine classification of h-scan parameters: assessment of steatosis in a rat model. *IEEE Trans Ultrason Ferroelectr Freq Control*. 2022;69(2):720-731.
21. Zhuang Z, Yang Z, Zhuang S, Joseph Raj AN, Yuan Y, Nersisson R. Multi-features-based automated breast tumor diagnosis using ultrasound image and support vector machine. *Comput Intell Neurosci*. 2021;2021:9980326.
22. He KM, Zhang XY, Ren SQ, Sun J, IEEE. Deep residual learning for image recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE; 2016:770-778.
23. Iandola FN, Han S, Moskewicz MW, Ashraf K, Dally WJ, Keutzer K. *SqueezeNet: AlexNet-Level Accuracy With 50x Fewer Parameters and <0.5MB Model Size*. arXiv:1602.07360. Accessed March 1, 2024. https://arxiv.org/abs/1602.07360
24. ImageNet large scale visual recognition competition (ILSVRC). Accessed March 1, 2024. http://www.image-net.org/

25. Kingma DP, Ba J. *Adam: A Method for Stochastic Optimization*. arXiv: 1412.6980. Accessed March 1, 2024. https://arxiv.org/abs/1412.6980
26. Sutskever I, Martens J, Dahl G, Hinton G. On the importance of initialization and momentum in deep learning. In: Sanjoy D, David M, eds. *Proceedings of the 30th International Conference on Machine Learning*. Vol 28. Proceedings of Machine Learning Research; 2013:1139-1147.
27. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. *Grad-CAM: Visual Explanations from Deep Networks Via Gradient-Based Localization*. arXiv:1610.02391. Accessed March 1, 2024. https://arxiv.org/abs/1610.02391
28. Tagaya R, Kurimoto N, Osada H, Kobayashi A. Automatic objective diagnosis of lymph nodal disease by B-mode images from convex-type echobronchoscopy. *Chest*. 2008;133(1):137-142.
29. Ozcelik N, Ozcelik AE, Bulbul Y, Oztuna F, Ozlu T. Can artificial intelligence distinguish between malignant and benign mediastinal lymph nodes using sonographic features on EBUS images? *Curr Med Res Opin*. 2020;36(12): 2019-2024.
30. Yong SH, Lee SH, Oh SI, et al. Malignant thoracic lymph node classification with deep convolutional neural networks on real-time endobronchial ultrasound (EBUS) images. *Transl Lung Cancer Res*. 2022;11(1):14-23.
31. Churchill IF, Gatti AA, Hylton DA, et al. An artificial intelligence algorithm to predict nodal metastasis in lung cancer. *Ann Thorac Surg*. 2022;114(1):248-256.
32. Ito Y, Nakajima T, Inage T, et al. Prediction of nodal metastasis in lung cancer using deep learning of endobronchial ultrasound images. *Cancers (Basel)*. 2022;14(14):3334.
33. Li J, Zhi X, Chen J, et al. Deep learning with convex probe endobronchial ultrasound multimodal imaging: a validated tool for automated intrathoracic lymph nodes diagnosis. *Endosc Ultrasound*. 2021;10(5):361-371.
34. Lindeman NI, Cagle PT, Aisner DL, et al. Updated molecular testing guideline for the selection of lung cancer patients for treatment with targeted tyrosine kinase inhibitors: guideline from the College of American Pathologists, the International Association for the Study of Lung Cancer, and the Association for Molecular Pathology. *Arch Pathol Lab Med*. 2018;142(3):321-346.
35. Singh N, Temin S, Baker S Jr, et al. Therapy for stage IV non-small-cell lung cancer with driver alterations: ASCO Living Guideline. *J Clin Oncol*. 2022;40(28): 3310-3322.
36. Singh N, Temin S, Baker S Jr, et al. Therapy for stage IV non-small-cell lung cancer without driver alterations: ASCO Living Guideline. *J Clin Oncol*. 2022; 40(28):3323-3343.
37. Vilmann P, Clementsen PF, Colella S, et al. Combined endobronchial and oesophageal endosonography for the diagnosis and staging of lung cancer. European Society of Gastrointestinal Endoscopy (ESGE) Guideline, in cooperation with the European Respiratory Society (ERS) and the European Society of Thoracic Surgeons (ESTS). *Eur Respir J*. 2015;46(1):40-60.
38. Asamura H, Suzuki K, Kondo H, Tsuchiya R. Where is the boundary between N1 and N2 stations in lung cancer? *Ann Thorac Surg*. 2000;70(6):1839-1845; discussion 1845-1836.
39. Osaki T, Nagashima A, Yoshimatsu T, Tashima Y, Yasumoto K. Survival and characteristics of lymph node involvement in patients with N1 non-small cell lung cancer. *Lung Cancer*. 2004;43(2):151-157.
40. Koseoglu FD, Alici IO, Er O. Machine learning approaches in the interpretation of endobronchial ultrasound images: a comparative analysis. *Surg Endosc*. 2023; 37(12):9339-9346.

Key Words: endobronchial ultrasound, artificial intelligence, deep learning, convolutional neural networks, diagnostic accuracy

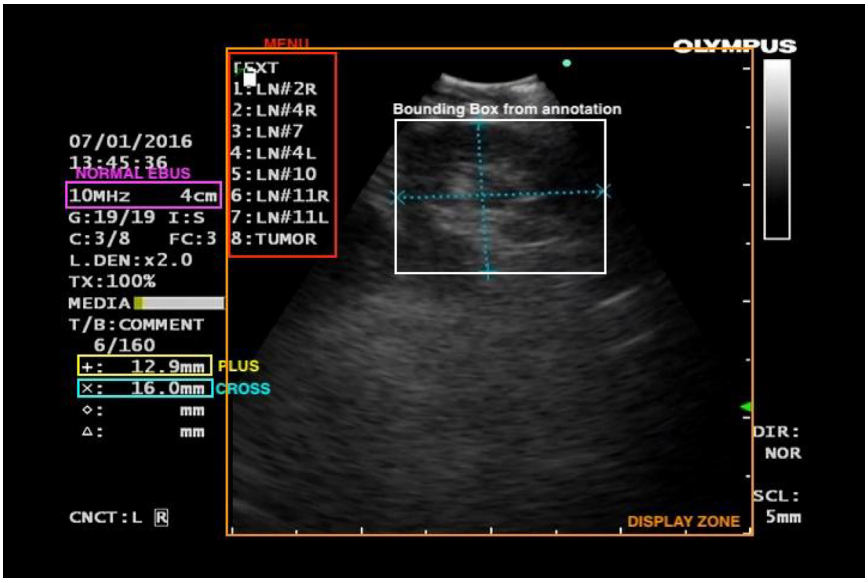


FIGURE E1. An example of an annotated video frame from the automated annotation tool. When the lymph node size is measured by the proceduralist, the annotation system marks it as PLUS (*yellow*) and CROSS (*turquoise*). When the labeling menu is opened to input the lymph node station name, the annotation system marks it as MENU (*red*).

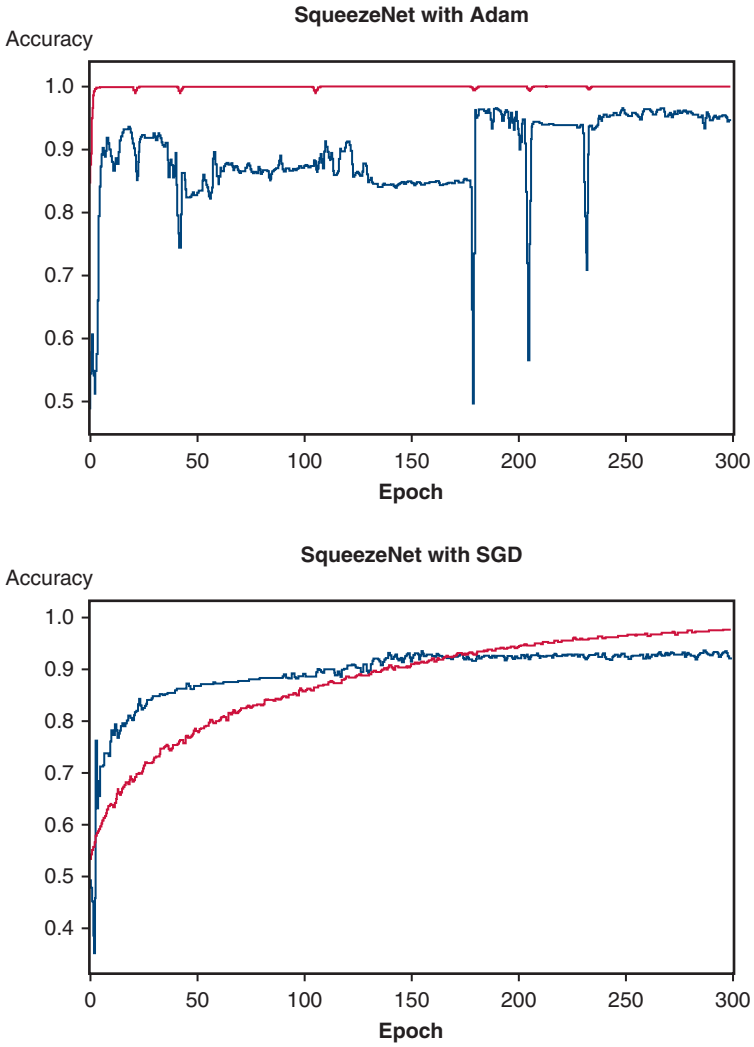


FIGURE E2. The performance curve for SqueezeNet architecture. *Red* curves represent training accuracy across epochs; *blue* curves represent validation accuracy. *Adam*, Adaptive moment estimation; *SGD*, stochastic gradient descent.

TABLE E1. Comparison of area under the receiver operating characteristic curves

Prediction model	AUC	Sensitivity* (%)	Specificity* (%)
Lymph node size measurement			
Short axis	0.86	75.6	82.2
Long axis	0.83	84.4	73.3
Long-to-short axis ratio	0.75	66.7	82.2
SVM (nondeep learning)	0.65		
SqueezeNet with Adam	0.99		
SqueezeNet with SGD	0.94		

AUC, Area under the curve; *SVM*, support-vector machine; *Adam*, adaptive moment estimation; *SGD*, stochastic gradient descent. *Sensitivity and specificity at the maximum Youden's index (sensitivity + specificity −1).