OPEN

PRS GLOBAL OPEN

ORIGINAL ARTICLE
Technology

Artificial Intelligence–Based Indocyanine Green Lymphography Pattern Classification for Management of Lymphatic Disease

Berk B. Ozmen, MD
Sonia K. Pandey, MD
Graham S. Schwarz, MD, MSE, FACS

Background: Lymphedema diagnosis relies on effective imaging of the lymphatic system. Indocyanine green (ICG) lymphography has become an essential diagnostic tool, but globally accepted protocols and objective analysis methods are lacking. In this study, we aimed to investigate artificial intelligence (AI), specifically convolutional neural networks, to categorize ICG lymphography images patterns into linear, reticular, splash, stardust, and diffuse.
Methods: A dataset composed of 68 ICG lymphography images was compiled and labeled according to five recognized pattern types: linear, reticular, splash, stardust, and diffuse. A convolutional neural network model, using MobileNetV2 and TensorFlow, was developed and coded in Python for pattern classification.
Results: The AI model achieved 97.78% accuracy and 0.0678 loss in categorizing images into five ICG lymphography patterns, demonstrating high potential for enhancing ICG lymphography interpretation. The high level of accuracy with a low loss achieved by our model demonstrates its effectiveness in pattern recognition with a high degree of precision.
Conclusions: This study demonstrates that AI models can accurately classify ICG lymphography patterns. AI can assist in standardizing and automating the interpretation of ICG lymphographic imaging. *(Plast Reconstr Surg Glob Open 2024; 12:e6132; doi: 10.1097/GOX.0000000000006132; Published online 23 August 2024.)*

INTRODUCTION

Indocyanine green (ICG) lymphography is a valuable diagnostic tool for the lymphatic system and the clinical management of lymphedema.[1] This imaging technique is recognized for its ability to visualize lymphatic vessels, provide critical insights into the lymphatic flow, and identify potential obstructions or anomalies.[2] The utilization of ICG, a fluorescent dye, in conjunction with a near-infrared spectral camera, allows for the real-time assessment of lymphatic functionality, marking a significant advancement over traditional lymphatic imaging methods, which often lack the resolution and specificity to visualize fine lymphatic structures.[3,4]

Despite its advantages, there is still not a globally accepted ICG lymphography protocol to conduct and interpret the results.[5] The interpretation of ICG lymphography remains subjective, relying on the expertise of clinicians. This subjectivity underscores the necessity for standardized and objective methods of analysis. The recent innovations in artificial intelligence (AI), particularly computer vision and convolutional neural networks (CNNs), present promising developments to address these limitations.[6,7] CNNs are well-suited for analyzing complex visual patterns in medical images due to their ability to automatically learn hierarchical features from raw pixel data.[8] Previous studies have successfully applied CNNs for various medical imaging tasks, such as diabetic retinopathy detection and skin cancer classification.[9,10] The ability of AI to learn from and interpret complex patterns in imaging could improve the analysis of ICG lymphography, providing reproducible and accurate methods of assessment.[11] In this study, we aimed to explore the efficacy of AI, specifically CNN algorithms, in classifying ICG lymphography patterns: linear, reticular, splash, stardust and diffuse.

From the Department of Plastic Surgery, Cleveland Clinic, Cleveland, Ohio.

Received for publication May 16, 2024; accepted July 17, 2024.

DOI: 10.1097/GOX.0000000000006132

METHODS

Patient Selection

This institutional review board approved study included ICG lymphography images collected retrospectively from patients who underwent ICG lymphography at a single institution. Patients were eligible for inclusion if

Disclosure statements are at the end of this article, following the correspondence information.

they were aged 18 or older with suspected or diagnosed lymphedema, and had no contraindications to ICG injection. Patients were excluded if they had a history of allergic reactions to iodine or indocyanine green, were pregnant, or had active infections in the extremity of interest.

ICG Lymphography Protocol

Patients were positioned supine on the examination table. Upper extremity studies were performed, with the arm resting on an armboard at a 60-degree angle to the body. Lower extremity images were captured with the leg extended on the procedure table. The near-infrared camera was positioned approximately 15–20 cm above the extremity, ensuring that the camera's field of view encompassed to optimize ICG signal intensity.

The ICG solution was prepared by combining one vial of ICG (25 mg) with 10 mL of saline. For upper extremity imaging, a total of 0.8 mL of the ICG solution was injected intradermally at four sites: the first and fourth web spaces of the hand and the volar wrist. For lower extremity imaging, a total of 0.8 mL was injected at four sites as described by Suami et al.[12] These sites maximize lymphosomal uptake of the upper and lower extremities for ICG lymphography imaging.

Immediately after the ICG solution was completely injected, a dynamic scan was performed to capture still images and video of the ICG fluorescence as it propagated through the lymphatic vessels. Following the dynamic scan, the patient was instructed to perform gentle exercises, such as hand clenching or foot pedaling, for 15 minutes. External interventions like milking and massaging were avoided to prevent interference with the natural flow dynamics of the ICG. Finally, a delayed scan was performed to capture still images and record additional videos.

Dataset Compilation and Labeling

A total of 68 ICG lymphography static images from the delayed scan were collected. The images were not categorized by upper or lower extremity, laterality, or specific anatomical locations and landmarks within the extremities. This is purposefully done to ensure the generalizability of the AI model for both upper and lower extremity lymphedema. These images were labeled by the study authors to identify five distinct lymphatic flow patterns: linear (Fig. 1), reticular (Fig. 2), splash (Fig. 3), stardust (Fig. 4), and diffuse (Fig. 5). To ensure label consistency, the manual labeling process was performed independently by two study authors experienced in recognizing ICG patterns. Any discrepancies in labeling were resolved through

Takeaways

Question: Can artificial intelligence accurately classify indocyanine green (ICG) lymphography patterns to enhance ICG lymphography procedures for diagnosis and treatment of lymphedema?

Findings: We developed a convolutional neural network model, which achieved 97.78% accuracy in categorizing ICG lymphography images into five recognized patterns: linear, reticular, splash, stardust, and diffuse. The model's high performance demonstrates its potential for automating and standardizing ICG lymphography interpretation.

Meaning: Artificial intelligence–based classification of ICG lymphography patterns can significantly enhance the accuracy and objectivity of lymphedema diagnosis, potentially improving clinical decision-making and patient outcomes.

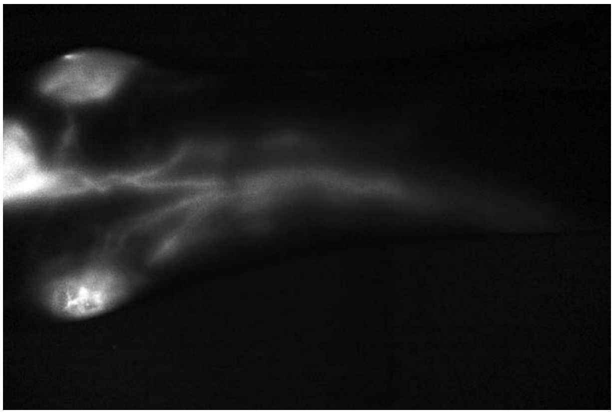

Fig. 1. Image showing the linear pattern.

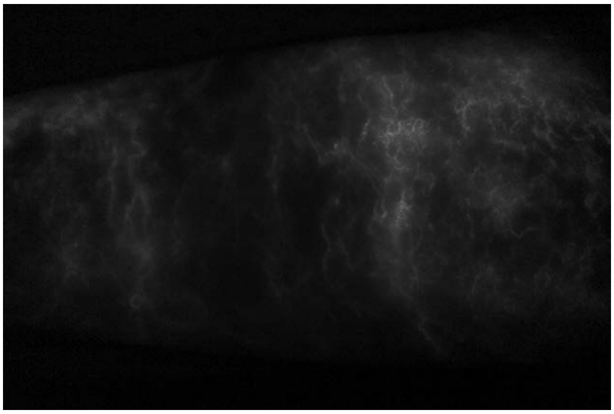

Fig. 2. Image showing the reticular pattern.

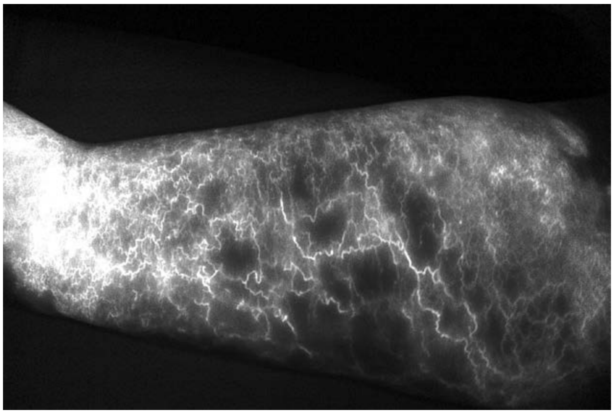

Fig. 3. Image showing the splash pattern.

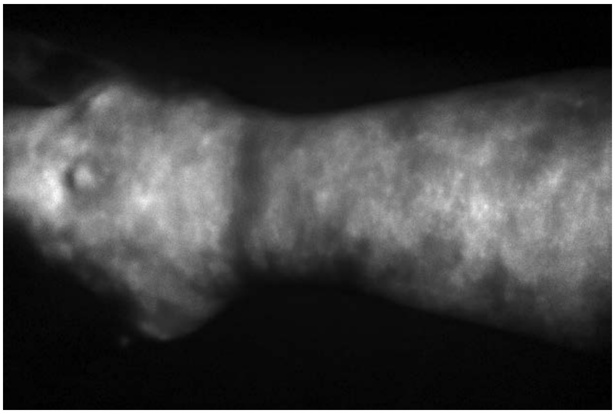

Fig. 4. Image showing the stardust pattern.

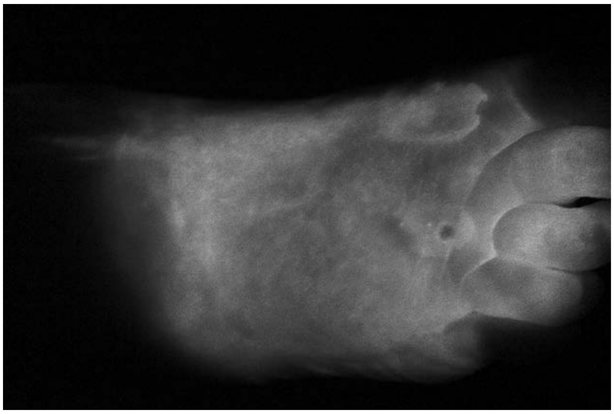

Fig. 5. Image showing the diffuse pattern.

consensus discussion with a third expert, the senior author of the study. The dataset included five linear, 15 reticular, seven splash, 21 stardust, and 20 diffuse patterns. Images were selected randomly but to reflect one pattern within one single frame. Images with multiple patterns in one single frame were not included in the dataset.

Dataset Preprocessing

Data were divided into three subsets: 80% for training, 10% for validation, and 10% for testing, ensuring a random distribution to mitigate bias and overfitting. The choice of an 80% training, 10% validation, and 10% testing data split was based on common practices in machine learning to ensure sufficient data for model training while allowing for performance evaluation on unseen data.[13] Before training, images underwent preprocessing in a computational environment to standardize their size and enhance contrast, optimizing them for neural network analysis.

AI Model Design and Training

We selected CNNs as an ideal algorithm for our study to evaluate ICG lymphography images due to their proven efficacy and accuracy in analyzing complex visual patterns in images.[14,15] Our model was developed using TensorFlow, an open-source platform for machine learning, and Python 3, known for its simplicity and efficiency in coding.[16,17]

We used the MobileNetV2 architecture as our base model, which is pretrained on the extensive ImageNet dataset.[18] This choice was given due to MobileNetV2's efficiency and effectiveness in mobile and embedded vision applications, along with its capability for transfer learning. Transfer learning allowed us to leverage the pretrained model's learned features, significantly reducing the need for extensive computational resources and training time for our study.[19] To design the MobileNetV2 model for our specific task of classifying ICG lymphography patterns, we used custom dense layers to the pretrained base. These layers were designed to refine the model's output to accurately reflect the five recognized lymphatic flow patterns. All coding and AI model development processes were carried out by the study authors.

Enhancing Model Performance with Data Augmentation

We implemented data augmentation techniques to bolster the model's ability to generalize across varied lymphography images.[20] These included random rotations, width and height shifts, shear transformations, zoom, and horizontal flipping. Data augmentation artificially expands the training dataset by generating transformed versions of the training images, thereby providing the model with a broader range of lymphatic flow patterns to learn from. This approach is instrumental in preventing overfitting and enhancing the model's robustness to variations in new, unseen images.[21]

AI Model Evaluation

The performance of our AI model was rigorously evaluated using the separate test set that we randomly partitioned in the beginning and kept aside. The primary metrics for evaluation were accuracy and loss, which were monitored throughout the training process. Accuracy metric measures the proportion of correctly predicted images out of the total, whereas loss metric quantifies the difference between the predicted patterns and the actual patterns, serving as an indicator of the model's error rate.[22] These metrics were pivotal in assessing the model's capability to classify ICG lymphography patterns accurately and reliably, providing a quantitative basis for the efficacy of integrating AI into lymphatic system imaging analysis.

RESULTS

Our CNN model demonstrated success in classifying ICG lymphography patterns, achieving a high accuracy rate of 97.78%. Throughout the training phase, which spanned 50 epochs, our model exhibited steady improvement in both accuracy and loss metrics. The accuracy metric indicates the model could correctly identify all five ICG lymphography patterns in test set images. The model's loss, quantifying the discrepancy between the predicted patterns and the actual patterns, was notably low at 0.0678. This low loss value signifies that the model predictions were remarkably close to the true classifications, further underscoring our model's reliability in analyzing ICG lymphography images.

The analysis of the model's performance across different lymphatic flow patterns revealed a consistent level of accuracy, with minimal variation in its ability to recognize the five predefined patterns: linear, reticular, splash, stardust, and diffuse. This uniformity in performance is indicative of the model's robust feature extraction and learning capabilities, enabling it to discern and categorize the complex visual characteristics inherent in each pattern type.

DISCUSSION

This study represents the very first research to incorporate AI into ICG lymphography and has successfully demonstrated that AI can be implemented to provide accurate and objective interpretation and classification of ICG lymphographic patterns. The high accuracy (97.78%) and low loss (0.0678) achieved by our model underscores the potential of AI and computer vision to enhance ICG lymphography. These findings suggest that AI could assist clinicians by standardizing the interpretation of lymphatic function, staging lymphedema severity, informing treatment decisions, and monitoring treatment response.

Our findings align with the broader literature that recognizes the value of computer vision and AI applications in plastic surgery.[6,23–28] Our model uses transfer learning to leverage the extensive data available in the ImageNet database, thereby reducing the need for large, labeled datasets which are often scarce in medicine.[29] The high accuracy rate achieved by our CNN model is indicative of its robustness in pattern recognition, even with a relatively small dataset. The application of data augmentation techniques contributed to the model's ability to generalize well to new data, thereby enhancing its predictive performance. These results are promising, especially considering the complex nature of ICG lymphography patterns and the importance of precise pattern classification in clinical practice.

Our model's capability to accurately predict the visual characteristics among linear, reticular, splash, stardust and diffuse patterns has significant clinical implications. Accurate discrimination between these specific patterns provides insight into the underlying physiology and functionality of lymph vessels and flow.[30] Linear patterns indicate normal lymph flow along a contractile vessel, while reticular and diffuse patterns suggest impairment.[31] Equipping clinicians with an objective, reproducible method to extract this level of detail empowers more informed clinical decision-making regarding disease staging, preoperative planning, and patient outcomes.

The consistency of the model across various images, without the need for specific considerations of laterality or anatomical landmarks, demonstrates its wide applicability. This suggests the model's potential for analyzing lymphatic patterns across different anatomical regions, supporting its use in diverse clinical settings. Another noteworthy potential application is use in tracking longitudinal changes of lymphatic patterns over serial assessments. Due to its high sensitivity, the model can detect subtle changes, enabling early detection of lymphatic dysfunction. This early detection capability allows opportunity for prompt intervention and halting of the progression of lymphedema.[32] Longitudinal monitoring of ICG patterns aided by this model could also help lend additional objectivity to outcome assessments of therapeutic and surgical interventions for lymphatic disorders.

The complex visual patterns in ICG lymphography studies often pose a steep learning curve for trainees in lymphatic imaging interpretation.[33] Integrating user-friendly AI tools with robust diagnostic performance as demonstrated here with our AI model could help flatten that learning curve and improve the clinical skills of surgical trainees and early-career lymphatic medicine practitioners.

Global adoption of ICG lymphography has been gradual despite its advantages over alternative imaging modalities.[34] Computer assisted review of this fluorescent guided imaging technique can provide an accessible and scalable solution to lessen the reliance on the availability of highly specialized on-site expertise for results interpretation.[35,36]

Explainability and interpretability of the AI model are crucial aspects to consider, especially in clinical applications. Understanding how the model makes its predictions can increase clinicians' trust and facilitate the integration of AI-assisted decision support into clinical workflows. Future work should explore techniques such as visualizing the model's attention maps to provide insights into the image regions that contribute to the model's predictions.[37]

Although our results demonstrate the efficacy of AI in enhancing ICG lymphography imaging, it is essential to acknowledge the limitations of this study. Our model was trained on a small dataset size of 68 images from a single institution. The dataset, though diverse and augmented, was limited to images from a single institution and single ICG lymphography device, which may affect the generalizability of the model. Future research should aim to include a larger and more varied dataset of lymphatic pathologies, potentially sourced from multiple institutions and devices, to further validate the model's efficacy. This study did not account for interobserver variability in the data labeling process. Future iterations of this research could incorporate a consensus approach or multiple expert validations to label the ICG lymphography patterns, enhancing the reliability of the training data. Although beyond the scope of the current study, comparing model performance to interpretations by a panel of expert clinicians could better quantify the advancements over traditional methods. Aggregate costs of image acquisition and processing, model development, and computer interface development warrant additional analysis.

CONCLUSIONS

This study demonstrates the significant potential of integrating AI into the analysis of ICG lymphography for automated classification of ICG lymphography patterns. By achieving a high accuracy rate of 97.78%, our AI model offers a promising solution to the challenges of subjectivity and variability among current clinical practices. Integration of AI-based decision support systems can help provide objective, reliable, and reproducible assessments of lymphatic function, improving the clinical and surgical management of lymphatic disease.

Graham S. Schwarz, MD, MSE, FACS
9500 Euclid Avenue, A60
Cleveland, OH
E-mail: schwarg@ccf.org

DISCLOSURE

The authors have no financial interests to declare in relation to the content of this article.

REFERENCES

1. Akita S, Unno N, Maegawa J, et al. A phase III, multicenter, single-arm study to assess the utility of indocyanine green fluorescent lymphography in the treatment of secondary lymphedema. *J Vasc Surg Venous Lymphat Disord*. 2022;10:728–737.e3.
2. Hara H, Mihara M. Multilymphosome injection indocyanine green lymphography can detect more lymphatic vessels than lymphoscintigraphy in lymphedematous limbs. *J Plast Reconstr Aesthet Surg*. 2020;73:1025–1030.
3. Mihara M, Hara H, Araki J, et al. Indocyanine green (ICG) lymphography is superior to lymphoscintigraphy for diagnostic imaging of early lymphedema of the upper limbs. *PLoS One*. 2012;7:e38182.
4. Nagy BI, Mohos B, Tzou CHJ. Imaging modalities for evaluating lymphedema. *Medicina (Kaunas)*. 2023;59:2016.
5. Sakai H, Miyazaki T. Need for better protocol and severity staging system of indocyanine green lymphography. *J Plast Reconstr Aesthet Surg*. 2023;79:19–20.
6. Nomura Y, Hoshiyama M, Akita S, et al. Computer-aided diagnosis for screening of lower extremity lymphedema in pelvic computed tomography images using deep learning. *Sci Rep*. 2023;13:16214.
7. Goudarzi S, Whyte J, Boily M, et al. Segmentation of arm ultrasound images in breast cancer-related lymphedema: a database and deep learning algorithm. *IEEE Trans Biomed Eng*. 2023;70:2552–2563.
8. Aggarwal R, Sounderajah V, Martin G, et al. Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. *NPJ Digit Med*. 2021;4:1–23.
9. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316:2402–2410.
10. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542:115–118.
11. Yamamoto T. Comprehensive lymphedema evaluation using dynamic ICG lymphography. In: Kusano M, Kokudo N, Toi M, Kaibori M, eds. *ICG Fluorescence Imaging and Navigation Surgery*. Japan: Springer; 2016:443–453.
12. Suami H, Thompson B, Mackie H, et al. A new indocyanine green fluorescence lymphography protocol for diagnostic assessment of lower limb lymphoedema. *J Plast Reconstr Aesthet Surg*. 2022;75:3946–3955.
13. Raschka S. Model evaluation, model selection, and algorithm selection in machine learning. Published online November 10, 2020.
14. Alzubaidi L, Zhang J, Humaidi AJ, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *J Big Data*. 2021;8:53.
15. Magnéli M, Borjali A, Takahashi E, et al. Application of deep learning for automated diagnosis and classification of hip dysplasia on plain radiographs. *BMC Musculoskelet Disord*. 2024;25:117.
16. Abadi M, Barham P, Chen J, et al. TensorFlow: a system for large-scale machine learning. In Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16), November 2–4, 2016, Savannah, GA. 2016:265–283. Available at https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi. Accessed February 12, 2024.
17. Abadi M, Agarwal A, Barham P, et al. TensorFlow: large-scale machine learning on heterogeneous distributed systems. Published online March 16, 2016.
18. Sandler M, Howard A, Zhu M, et al. MobileNetV2: inverted residuals and linear bottlenecks. Published online March 21, 2019.
19. Falou O, Sannachi L, Haque M, et al. Transfer learning of pretreatment quantitative ultrasound multi-parametric images for the prediction of breast cancer response to neoadjuvant chemotherapy. *Sci Rep*. 2024;14:2340.
20. Chlap P, Min H, Vandenberg N, et al. A review of medical image data augmentation techniques for deep learning applications. *J Med Imaging Radiat Oncol*. 2021;65:545–563.
21. Chilukoti SV, Shan L, Tida VS, et al. A reliable diabetic retinopathy grading via transfer learning and ensemble learning with quadratic weighted kappa metric. *BMC Med Inform Decis Mak*. 2024;24:37.
22. Schlosser T, Friedrich M, Meyer T, et al. A consolidated overview of evaluation and performance metrics for machine learning and computer vision. 2023.
23. Knoedler L, Alfertshofer M, Simon S, et al. Diagnosing lagophthalmos using artificial intelligence. *Sci Rep*. 2023;13:21657.
24. Yu S, Dwight J, Siska RC, et al. Feasibility of intraoperative image guidance in burn excision surgery with multispectral imaging and deep learning. *Burns*. 2024;50:115–122.
25. Han SS, Moon IJ, Lim W, et al. Keratinocytic skin cancer detection on the face using region-based convolutional neural network. *JAMA Dermatol*. 2020;156:29–37.
26. Tokgöz E, Carro MA. Applications of artificial intelligence, machine learning, and deep learning on facial plastic surgeries. In: Tokgöz E, Carro MA, eds. *Cosmetic and Reconstructive Facial Plastic Surgery: a Review of Medical and Biomedical Engineering and Science Concepts*. Switzerland: Springer Nature; 2023:281–306.
27. Kenig N, Monton Echeverria J, De la Ossa L. Identification of key breast features using a neural network: applications of machine learning in the clinical setting of plastic surgery. *Plast Reconstr Surg*. 2024;153:273e–280e.
28. Khedgaonkar R, Singh K, Raghuwanshi M. Chapter 10. Local plastic surgery-based face recognition using convolutional neural networks. In: Pradeep N, Kautish S, Peng SL, eds. *Demystifying Big Data, Machine Learning, and Deep Learning for Healthcare Analytics*. Academic Press; 2021:215–246.
29. Deng J, Dong W, Socher R, et al. ImageNet: a large-scale hierarchical image database. Paper presented at: 2009 IEEE Conference on Computer Vision and Pattern Recognition; Miami, Fla. Available at https://ieeexplore.ieee.org/document/5206848. Accessed February 19, 2024.
30. Narushima M, Yamamoto T, Ogata F, et al. Indocyanine green lymphography findings in limb lymphedema. *J Reconstr Microsurg*. 2016;32:72–79.
31. van Heumen S, Riksen JJM, Bramer WM, et al. Imaging of the lymphatic vessels for surgical planning: a systematic review. *Ann Surg Oncol*. 2023;30:462–479.
32. Akita S, Nakamura R, Yamamoto N, et al. Early detection of lymphatic disorder and treatment for lymphedema following breast cancer. *Plast Reconstr Surg*. 2016;138:192e–202e.
33. Dip F, Alexandru N, Amore M, et al. Use of fluorescence imaging during lymphatic surgery: a Delphi survey of experts worldwide. *Surgery*. 2022;172:S14–S20.
34. Kim HB, Jung SS, Cho MJ, et al. Comparative analysis of preoperative high frequency color Doppler ultrasound versus MR lymphangiography versus ICG lymphography of lymphatic vessels in lymphovenous anastomosis. *J Reconstr Microsurg*. 2023;39:92–101.
35. Bekbolatova M, Mayer J, Ong CW, et al. Transformative potential of AI in healthcare: definitions, applications, and navigating the ethical landscape and public perspectives. *Healthcare (Basel)*. 2024;12:125.
36. Khanna NN, Maindarkar MA, Viswanathan V, et al. Economics of artificial intelligence in healthcare: diagnosis vs. treatment. *Healthcare (Basel)*. 2022;10:2493.
37. Selvaraju RR, Cogswell M, Das A, et al. Grad-CAM: visual explanations from deep networks via gradient-based localization. Paper presented at: 2017 IEEE International Conference on Computer Vision (ICCV) Venice, Italy, October 22-29, 2017; 2017:618–626.