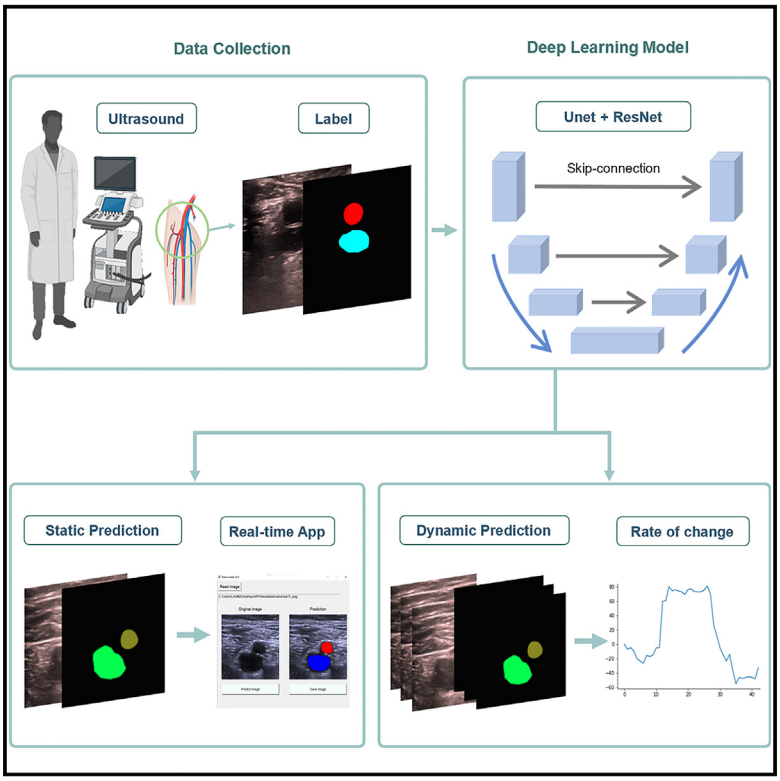
iScience

Article

Deep learning model for diagnosis of venous thrombosis from lower extremity peripheral ultrasound imaging

Graphical abstract



Authors

Po-Wei Chen, Bor-Yann Tseng, Zhu-Han Yang, Chi-Hua Yu, Keng-Tse Lin, Jhen-Nong Chen, Ping-Yen Liu

Correspondence

jonnyyu@gs.ncku.edu.tw (C.-H.Y.), larry@mail.ncku.edu.tw (P.-Y.L.)

In brief

Cardiovascular medicine; Machine learning; Ultrasound technology

Highlights

- The utility of UNet and ResNet architectures demonstrates enhanced image processing
- ResUNet enables automated segmentation of vessel walls from venous duplex ultrasound videos
- ResUNet outperforms UNet in vessel identification, offering potential for real-time analysis

Chen et al., 2024, iScience 27, 111318
December 20, 2024
https://doi.org/10.1016/j.isci.2024.111318

CellPress

iScience

CellPress
OPEN ACCESS

Article

Deep learning model for diagnosis of venous thrombosis from lower extremity peripheral ultrasound imaging

Po-Wei Chen,[1,2] Bor-Yann Tseng,[3] Zhu-Han Yang,[3] Chi-Hua Yu,[3,4,*] Keng-Tse Lin,[3] Jhen-Nong Chen,[3] and Ping-Yen Liu[1,2,*]

[1]Institute of Clinical Medicine, College of Medicine, National Cheng Kung University, Tainan 704, Taiwan
[2]Division of Cardiology, Department of Internal Medicine, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Tainan 704, Taiwan
[3]Department of Engineering Science, National Cheng Kung University, No. 1, University Road, Tainan 701, Taiwan
[4]Lead contact
*Correspondence: jonnyyu@gs.ncku.edu.tw (C.-H.Y.), larry@mail.ncku.edu.tw (P.-Y.L.)
https://doi.org/10.1016/j.isci.2024.111318

SUMMARY

Deep vein thrombosis (DVT) causes significant healthcare burdens worldwide. This study aims to establish a deep learning model for the diagnosis of DVT from the assessment of vein compressibility. Considering the complexity of ultrasound images, convolutional neural networks with UNet and residual neural network (ResNet) are established for image segmentation, from venous duplex ultrasonographic video images, obtained through standard and portable handheld ultrasound methods. To further evaluate the similarity between the predicted and ground truth images, the structural similarity index (SSIM) is employed. Our deep learning model achieves over 90% accuracy, providing an innovative tool for both images and videos. This study harnesses the power of machine learning to develop an automatic labeling tool that can diagnose DVT by analyzing ultrasonography images. To make the tool more accessible to front-line clinicians, a user-friendly application is created to quickly assess possible clinical severity and enable prompt medical intervention, reducing disease progression.

INTRODUCTION

Deep venous thrombosis (DVT) and pulmonary embolism (PE) are two clinical manifestations of venous thromboembolism (VTE). Untreated lower extremity DVT may cause fatal PE but overdiagnosis may increase the risk of anticoagulation. Many patients with similar symptoms do not actually have DVT, resulting in long referral times for patients and an increased clinical burden for specialists. Therefore, timely and accurate diagnosis of DVT is essential for clinical physicians.[1–3]

Despite many advances in imaging modalities, peripheral venous ultrasonography is one of the most prevalent diagnostic tools for the diagnosis of DVT.[1,4,5] Contrast venography, considered the reference modality for diagnosing DVT, is typically reserved for situations in which other tests fail to conclusively establish or exclude DVT. However, it is essential to underscore that compression ultrasonography should be regarded as the gold standard first-line imaging method for the majority of patients with an elevated likelihood of DVT[2,6–8] Thus, a diagnosis at the point of care by non-specialists is desired.

A key step in peripheral venous ultrasonography is the differentiation of veins from arteries. Veins are normally fully compressible under light pressure; in contrast, arteries require substantial pressure for compression. However, the manual compression process lacks a standard operating procedure.[8,9]

Clinical evidence has suggested that peripheral venous duplex ultrasonography can be performed by nurses or technicians.[5,10,11] However, confidence in acquiring basic images is generally low because of the required image interpretation skills and lack of standard protocol. Thus, the operation method is not widely applied in practice.[2,9]

The advent of artificial intelligence (AI) has revolutionized many sectors, with particularly profound implications in medical diagnostics.[12–14] Modern AI models can predict diverse targets, be they numerical, multi-class, or structured, and have made remarkable strides in areas such as cancer detection, neurological disease diagnosis, and cardiovascular assessments through cardiac ultrasonography and electrocardiography.[15–17] However, their optimal performance often hinges on the availability of substantial, high-quality, professionally labeled data. In the context of disease detection, this prerequisite poses challenges, given the professional expertise required for data labeling and the inherent difficulties in sourcing vast datasets.[18] As a countermeasure, the UNet was introduced primarily for cell edge segmentation. Its unique "U" shaped design efficiently processes medical images even when datasets are limited, making it invaluable in medical image segmentation applications.[19–21] However,

as AI models become more intricate, they risk reduced accuracy, especially with increased depth.[22,23] Recognizing this, the ResNet model emerged in 2015, providing a robust solution for various image processing tasks, from segmentation to recognition.[23] Its practicality and effectiveness have made it indispensable in diverse projects, underlining its transformative potential in AI-driven diagnostics.

To ascertain if deep learning technology can offer consistent image acquisition guidance and immediate diagnostic assistance, we probed its applicability in the diagnosis of VTE.[21] Current VTE diagnostic models predominantly lean on clinical data, such as the Wells criteria. While certain studies have suggested deep learning methodologies for interpreting ultrasound images to distinguish between DVT and non-DVT, none, to our knowledge, have delved into the advantages of UNet and ResNet in the image-centric diagnosis of venous hypertension.[6,24–27] Additionally, there's a discernible research void concerning the relationship between image-steered deep learning technology and clinical efficacy. Our investigation seeks to bridge these knowledge chasms, illuminating the potential of deep learning within this domain.

Automated interpretation of venous duplex images through deep neural networks represents a promising advancement for refining clinical documentation and enhancing operational efficiency. Our objective was to design and validate a deep neural network model tailored for video analysis, synergizing the strengths of UNet and ResNet blocks to facilitate automated vessel wall detection and compressibility assessment. This automation not only alleviates the workload on professionals responsible for post-ultrasound image annotation but also provides auxiliary assessments, enabling real-time and dynamic labeling as ultrasound images evolve during examinations.

To fulfill this vision, we fashioned a deep learning model endowed with swift predictive capabilities. It offers real-time labeling of arteries and veins in the provided images, substantially reducing dependence on manual labeling—a process acknowledged for being laborious and time-intensive. The model's commendable accuracy and versatility ensure robust performance, even with images that may be subpar in quality or significantly differ from those in the training dataset. When integrated with portable devices, our model provides immediate predictions, further enhancing clinical diagnostic processes and streamlining patient care.

In our pursuit to gauge the model's efficacy, we employed a dataset of venous duplex images and undertook the essential labeling tasks. This curated data informed our model's training phase, capacitating it to predict arterial and venous locations in static images and to render real-time predictions through a Graphical User Interface (GUI). The model's dynamic predictions also granted us the ability to quantify venous changes during compression, adding depth to our diagnostic insights.

The manuscript is structured as follows: We initiate by delineating the data acquisition methodologies and sources, then delve into data preprocessing and shed light on the model's architecture and its training process. Subsequently, we showcase the results achieved with our model on both static and dynamic images, taken before and after compression. Moreover, we benchmark our model's performance against the UNet architecture to affirm our method's robustness. In addition, we also ensure the predictive ability of the model for images from different sources through external verification. Collectively, our study illuminates the game-changing potential of deep neural networks in the realm of venous duplex interpretation and emphasizes the benefits of melding UNet and ResNet for achieving unparalleled precision and efficiency.

RESULTS

Model prediction accuracy

We utilized a ResUNet-based model to identify the location and area of blood vessels in ultrasound images for disease diagnosis assistance. To enhance training performance, we applied data augmentation to the dataset. The training process is summarized in Figure 1, where we used BCE LogDice Loss as the loss function to evaluate the model's prediction performance. As shown in Figure 1A, training and validation dataset has a loss equal to about 0.05 and an accuracy higher than 99%. The Dice coefficient was used to assess the model's image prediction accuracy and the training curve is depicted in Figure 1B. Dice coefficient between the test and the validation dataset is approximately equal to 0.98.

To further evaluate the similarity between the predicted and ground truth images, we employed the structural similarity index (SSIM), and the evaluation results are presented in Figure 1C. The results indicate that most of the arterial and venous structures had an SSIM value greater than 0.98. These findings demonstrate that our ResUNet-based model can accurately predict the location and area of blood vessels in ultrasound images, providing valuable assistance in disease diagnosis.

Comparison of deep learning models

In this study, we partitioned the dataset into distinct sets for training, validation, and testing, constituting 75%, 15%, and 10% of the total data, respectively. The training of the model was conducted on a GeForce RTX 2080Ti GPU, utilizing a batch size of 32 over the course of 5000 epochs. For the optimization process, the Adam optimizer was deployed, alongside a hybrid loss function integrating Binary Cross-Entropy with Dice Loss. This combination facilitated the model's enhanced capability in discerning both macroscopic and microscopic anatomical structures within ultrasound images. The meticulous design of our training and testing methodologies has been instrumental in the establishment of a model characterized by high precision and efficiency. This model is particularly adept at delineating the locations of blood vessels in ultrasonographic scans, thereby demonstrating its potential utility in medical imaging diagnostics. Figure 1D presents a comparative analysis of ResUNet and UNet in the automatic labeling task. The results indicate that the ResUNet model significantly outperforms UNet in identifying arteries and veins. Under consistent training conditions, ResUNet accurately predicted both artery and vein positions, while UNet predominantly approximated vein positions. The advanced architecture of ResUNet yielded more precise edge predictions, thereby enhancing segmentation accuracy. Furthermore, the magnified view highlights a significant failure of UNet in accurately segmenting veins. Quantitative analysis of the

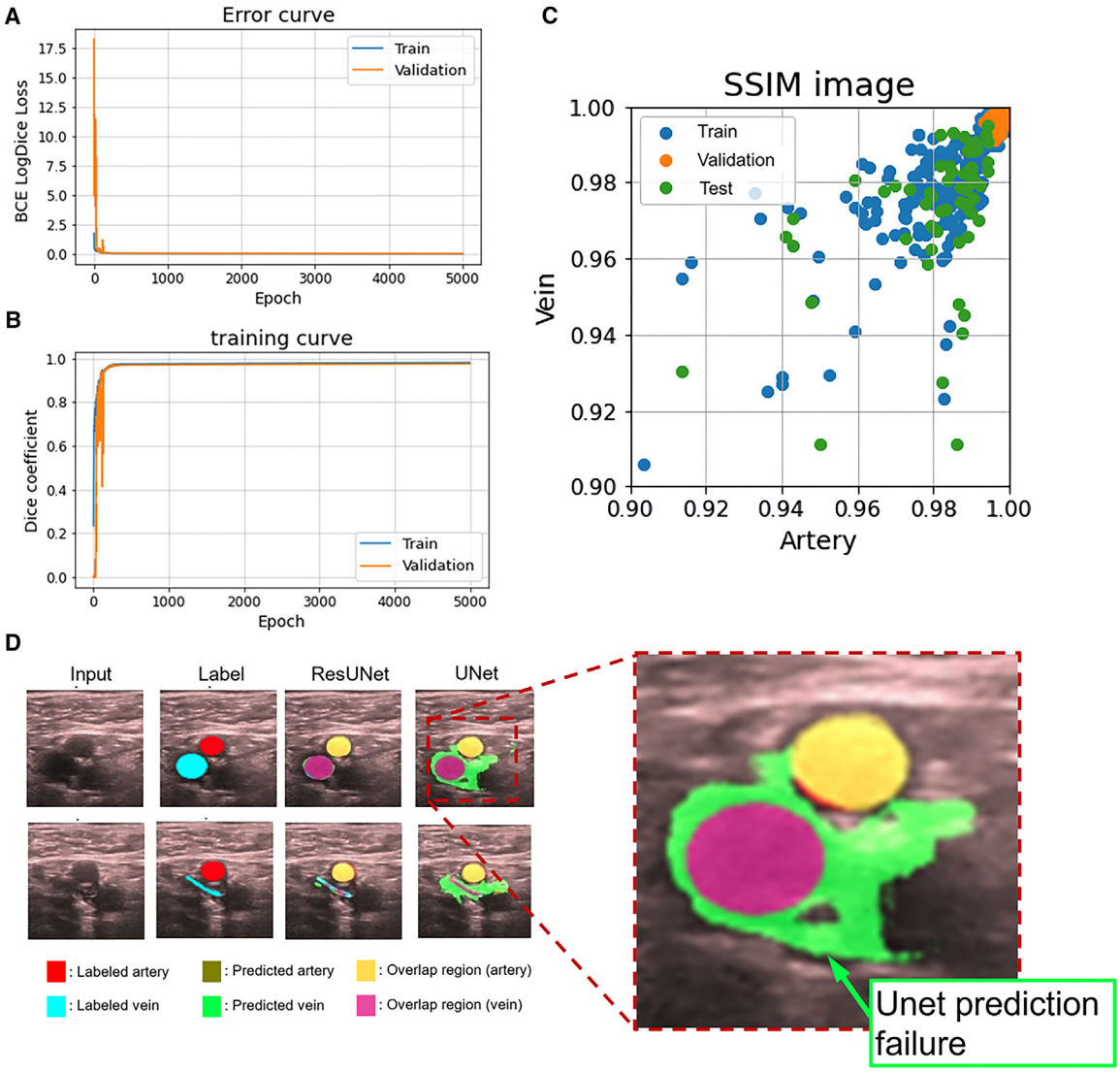


Figure 1. Training performance and image similarity

(A) The error curve indicates that the training dataset achieved a loss of approximately 0.05 with an accuracy exceeding 99%, whereas the validation dataset had a loss of around 0.06 with an accuracy exceeding 99%.

(B) The error curve demonstrates that both the training dataset and validation dataset achieved a Dice coefficient of approximately 0.98, indicating a high prediction accuracy of the model.

(C) Using SSIM (structural similarity) as an indicator of label and prediction image similarity, it can be seen that most of the images have a similarity greater than 0.98.

(D) Comparison of the performance of ResUNet and UNet at image segmentation and verification that the optimization of UNet was effective. The magnified view highlights a failure case of UNet in accurately segmenting the vein, effectively resolved by the ResUNet model.

misprediction areas revealed that UNet's prediction failure area was approximately 21 times greater than that of ResUNet. These findings underscore the superior segmentation accuracy of the ResUNet model.

To compare the segmentation results under the Dice and Jaccard coefficients. We employed the same settings used in the original model training, such as parameter configurations and datasets. We retrained the model and recorded both Dice and Jaccard coefficients throughout the training process. In Figure 2A, both the Dice and Jaccard coefficients experienced a rapid ascent in the initial training phases, demonstrating the model's quick adaptability to the segmentation tasks. The Dice coefficient stabilized near a high of 0.9, while the Jaccard coefficient plateaued around 0.8. This indicates that our model achieves high segmentation accuracy and maintains this performance consistently across epochs. We selected the same case to illustrate the prediction results under two different metrics. In Figures 2B and 2C, the attached image shows the segmentation results using both the Dice coefficient and the Jaccard coefficient. This comparison demonstrates that our model

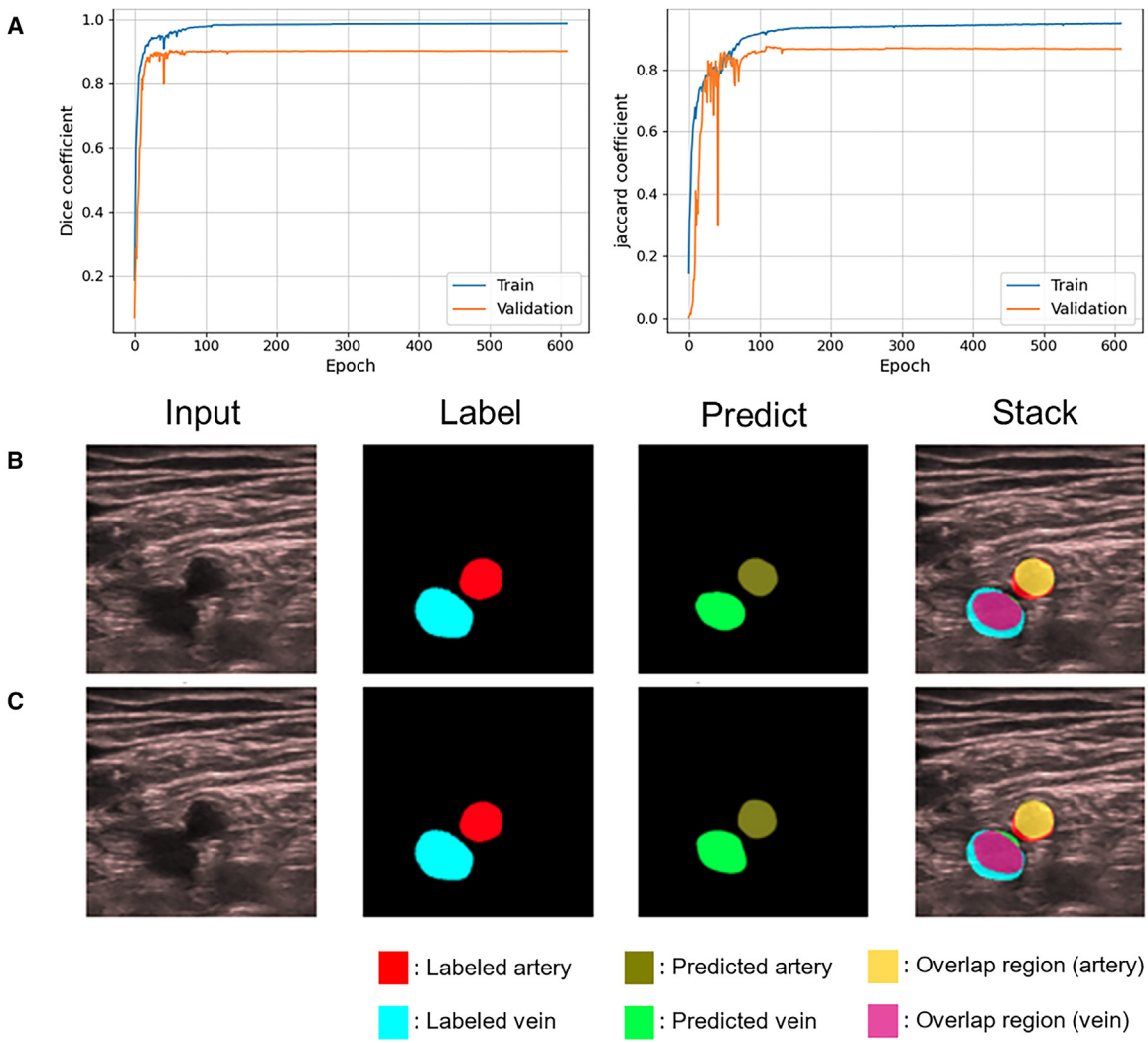


Figure 2. Training performance comparison and segmentation results using Dice and Jaccard coefficients

(A) The Dice and Jaccard coefficients training performance. The segmentation results using both the (B) Dice coefficient and the (C) Jaccard coefficient.

consistently performs well regardless of the metric used. These results highlight the model's excellent training performance and segmentation accuracy using both evaluation methods.

Static segmentation

In the corpus of ultrasound imagery utilized for this investigation, two primary classifications emerge: precompression and compression images. Within the precompression category, the delineation of arterial and venous structures is prominent, whereas the compression images display a marked diminution in venous caliber attributable to the applied pressure.

Employing the ResUNet-based architecture, our study has yielded promising results in the automated localization of vascular structures within both precompression and compression images. The veracity of the model's predictions is meticulously documented in Figure 3, providing compelling visual evidence of its ability to accurately render the positions of blood vessels across the two states of compression.

For a more rigorous validation of the model's localization precision, an overlay comparison was conducted between the labeled and algorithmically predicted images. Figure 3A illustrates this comparison, showing a substantial overlap that signifies a high degree of concordance in the model's predictions.

Further quantitative analysis was pursued through the calculation of pixel intensities associated with vascular structures in both image states. The labeled precompression images presented a pixel intensity sum of 1149, whereas the predicted images exhibited a sum of 1423. In the compression paradigm, these sums were 169 and 193, respectively. When assessing the relative change, the labeled images manifested a rate of 82.29%, and the predicted images demonstrated a rate of 86.44%. This parallelism in rate change further substantiates the predictive accuracy of the model.

The implications of these findings are noteworthy in the clinical context. The model's robust capacity for precise vascular localization in static ultrasound images—both under precompression

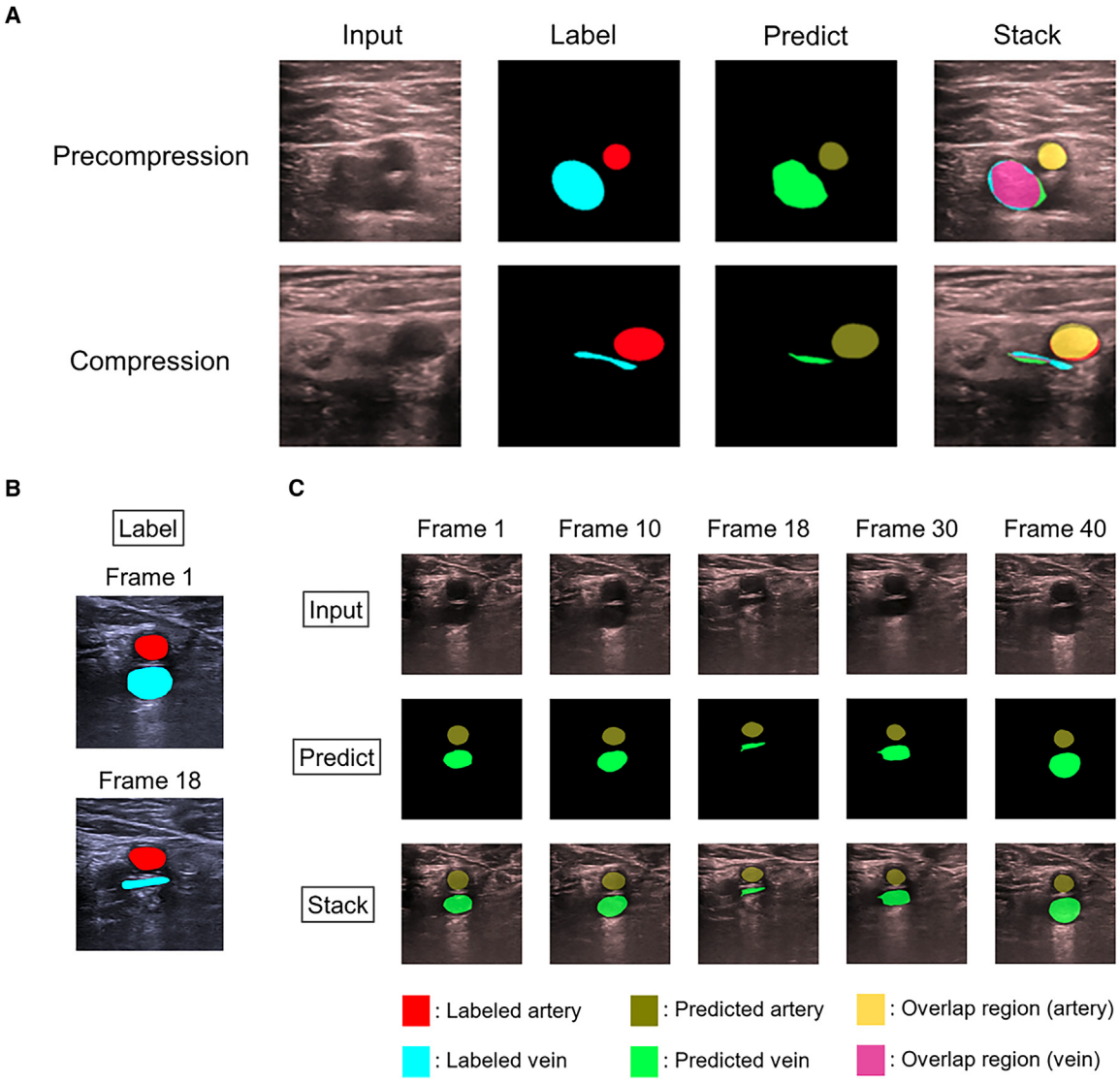


Figure 3. Static image and video prediction example

(A) The deep learning model accurately predicts the position of the artery and vein on the ultrasound images. Confirmation of training performance by stacking the labeled and predicted images.

(B) Frame of the labeled image.

(C) The deep learning model accurately predicts the position of the artery and vein in each frame of the ultrasound video. By comparing the same frame, the labeled and prediction results are similar.

and compression—serves as a pivotal tool for healthcare professionals. It aids in the expeditious identification of vein compressibility and supports the accurate diagnosis of vascular conditions. This utility is particularly salient in clinical scenarios where time is of the essence, offering a rapid, non-invasive diagnostic reference for the assessment of venous pressure and vessel location, obviating the need for manual interpretive analysis.

Dynamic segmentation

In the realm of medical imaging, the dynamic prediction of blood vessel location and area in ultrasound videos represents a significant advancement. Our research is primarily concentrated on this aspect, recognizing that a substantial portion of the images in the dataset originates from video sequences. This focus on video data necessitates a model capable of temporal and spatial precision.

Our methodology involved the segmentation of each video into static frames at an interval of every two frames. The model was then tasked with predicting the location of blood vessels in each isolated frame. This meticulous frame-by-frame analysis facilitated a granular prediction of arterial and venous positions throughout the video sequence.

Subsequent to the prediction phase, the frames were reconstituted into a continuous video format. This reconstruction allowed

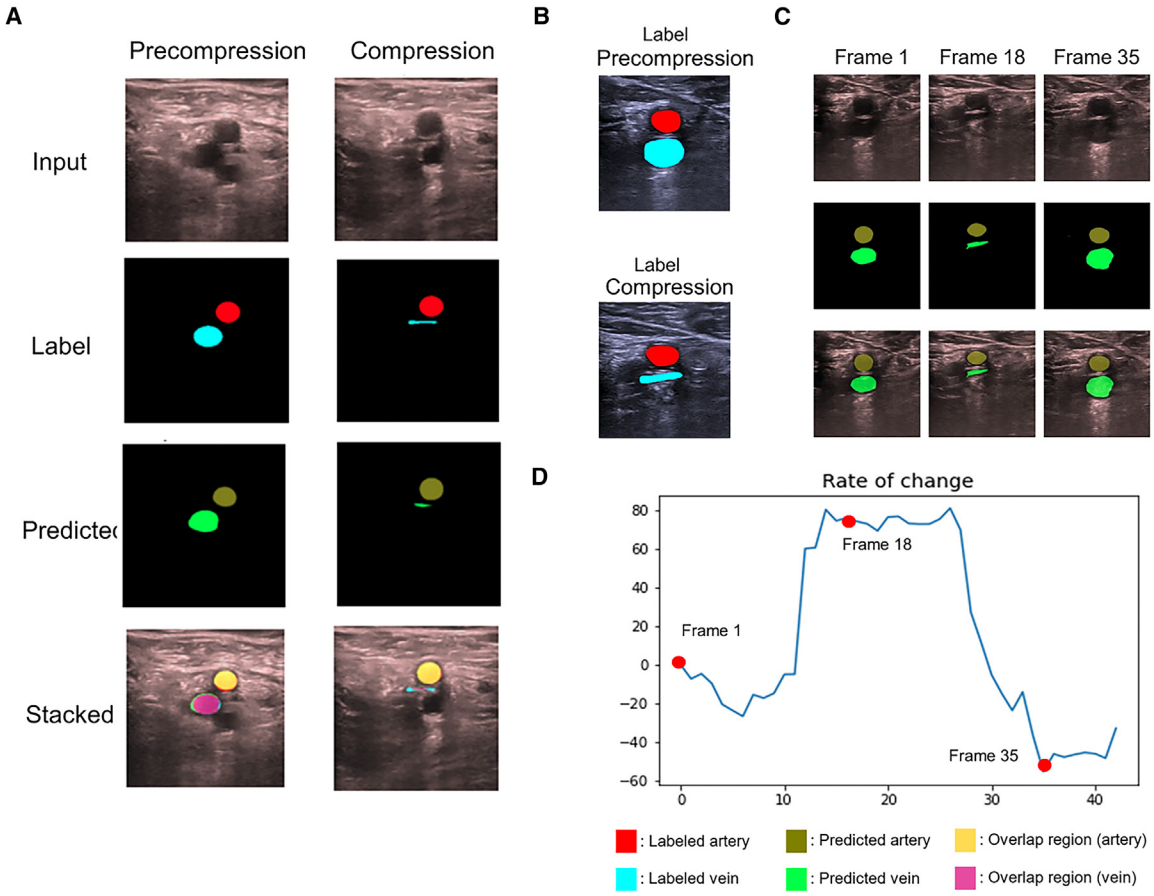


Figure 4. Static image and video prediction example of calculated rates of change

(A) For the precompression state of the blood vessel, we calculated the pixel quality of the labeled image to be 1149 and the pixel quality of the predicted image to be 1423. For the compression state of the blood vessel, we calculated the pixel quality of the labeled image to be 169 and the pixel quality of the predicted image to be 193. The rate of change of the labeled image was 82.29%, whereas that of the predicted image was 86.44%, demonstrating that the rates of change were similar.

(B) Labeled image precompression and after compression.

(C) Dynamic prediction result.

(D) Rate of change curve during the compression process calculated based on the vein pixel value of the first frame. The red dots are the corresponding frame of the image above the rate of change curve.

for an observation of the vascular morphology and its transformation during the process of compression, providing a comprehensive visualization of the dynamic changes occurring within the arterial and venous structures.

The proficiency of our model in dynamic prediction is demonstrable in Figures 3B and 3C. In Figure 3B, we exhibit the congruence between the labeled frames and those predicted by the model, establishing a benchmark for accuracy in video frame correspondence. Furthermore, Figure 3C illustrates the segmentation of a 2-s video into approximately 40 frames, enabling an in-depth visual examination of vascular dynamics during the compression phase.

The predictive outcomes for specific frames, such as Frame 1 and Frame 18 in Figure 3C, show a remarkable alignment with the professional annotations presented in Figure 3B. This correlation serves as an affirmation of the high precision of our model in the context of dynamic image prediction. The video selected for the demonstration of dynamic prediction was meticulously chosen from the test set, ensuring an unbiased assessment of the model's performance.

Degree of compressibility

In the nuanced field of vascular ultrasound diagnostics, the precise measurement of venous cross-sectional area changes holds significant clinical relevance, particularly in the diagnosis and monitoring of thrombotic disorders. This study introduces a methodological innovation, leveraging pixel intensity quantification to determine venous cross-sectional area from ultrasound images, with the results expounded in Figure 4.

Our investigation employed this advanced image analysis technique to both annotated and algorithmically predicted static images, thereby evaluating the accuracy of our computational model in estimating venous dimensions. As evidenced in Figure 4A, the rate of change in the venous cross-sectional

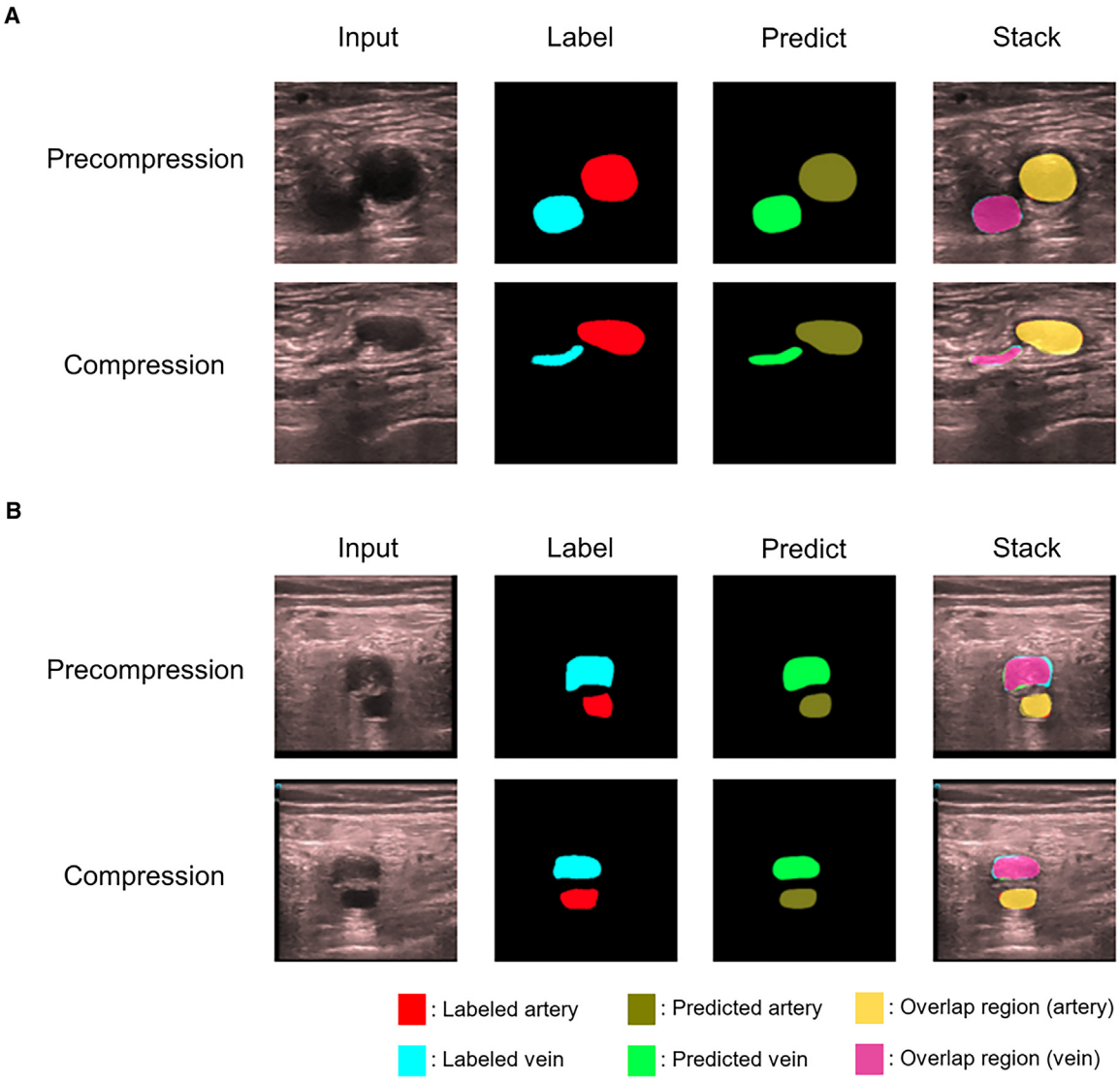


Figure 5. Comparison of vein compressibility in healthy control and DVT patient

(A) Ultrasound images of a normal femoral vein. After compression, the vein is completely collapsed, indicating normal compressibility.

(B) Ultrasound images from a patient of DVT. The vein does not collapse but has an oval shape indicating an acute DVT based on the noncompressible but deformable vein. Our ResUnet model can recognize the artery and vein, and measure vein compressibility to differentiate the presence of DVT.

area, calculated from the pixel values between precompression and compression images compared against labeled data, corroborates the model's capacity for precise prediction of venous deformation. This correlation is critical, as it provides a quantifiable measure of venous elasticity and patency, which are crucial determinants in the clinical assessment of thrombosis.

Conventionally, sonographic quantification of venous structures has been susceptible to operator-dependent variability, potentially impacting the consistency of thrombotic diagnoses. By incorporating our model into clinical practice, health care providers are equipped with a tool that offers an expedited, objective analysis of venous compression. This enhancement can lead to a more standardized and reliable approach to thrombosis detection and characterization.

Extending beyond static image analysis, our methodology was adeptly applied to dynamic video data to calculate the rate of change in venous cross-sectional area under various compressive states, showcased in Figures 4B and 4C. The dynamic video analysis facilitated a frame-by-frame quantification, with Frame 18 presenting a significant contraction ratio indicative of maximum compression. This data is further distilled into a rate of change curve during the compression cycle, as illustrated in Figure 4D, based on the vein pixel value in the initial frame of the video sequence. This curve provides a temporal map of the changing venous landscape, yielding

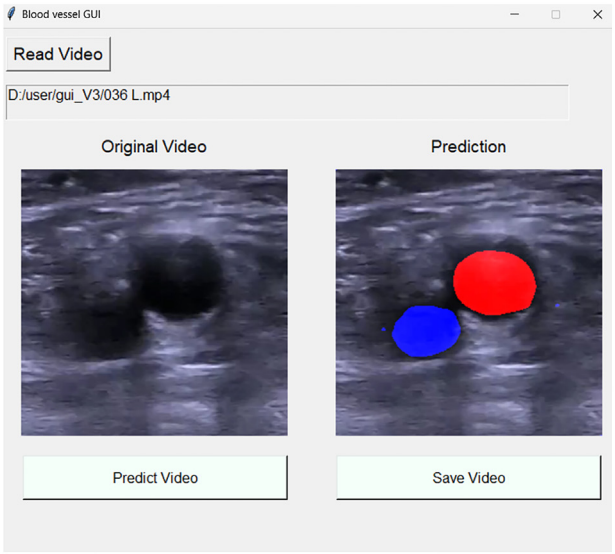


Figure 6. Graphical user interface

The GUI interface first allows doctors to import ultrasound videos and recognize the positions of the artery and vein. From these dynamic images, it provides the changes in the vein wall before and after compression for clinical reference. Finally, the GUI can save the predicted ultrasound video.

critical diagnostic information regarding thrombotic risk and aiding clinicians in the rapid identification and assessment of venous disorders.

Loss of compressibility of the vein is the most reliable indicator of the presence of thrombus within the vein. In individuals without DVT, compression by an ultrasound probe naturally flattens the vein. To illustrate these differences, two images are presented: one from a control group (individuals without DVT) and another from a patient with DVT. In Figure 5A, the control group image demonstrates significant vein deformation under compression, indicative of an unobstructed vein. In Figure 5B, the image from the DVT patient shows decreased compressibility, suggesting elevated venous pressure and indicating possible obstruction by a blood clot.

To bridge the gap between complex computational models and clinical utility, our team has engineered a graphical user interface (GUI), leveraging the Tkinter library, to democratize access to our vascular positioning model. The interface, delineated in Figure 6, is designed with an emphasis on ease of use, allowing clinicians to rapidly ascertain the locational data of arterial and venous structures within ultrasound images without the necessity for programming acumen.

Upon engaging with the GUI, users are presented with an option to upload the ultrasound video of interest. The interface provides real-time visualization of the predicted outcomes, rendering a visual prediction of vascular positions. Users can then save the results as GIF files via the 'Save' feature. This sequence of operations is streamlined to ensure the GUI's alignment with the time-sensitive demands of clinical practice. Additionally, we have created a tutorial video demonstrating the interface's capabilities. For those interested, a QR code and a URL (https://youtu.be/wMxeMsQN28E) are provided to access this video, offering a convenient way to see the new functionality in action.

The rationale for the GUI's development was to encapsulate the sophistication of the underlying predictive model within an accessible, user-friendly platform, thereby mitigating potential barriers to its adoption in clinical settings. The interface serves as a conduit, translating advanced image processing into actionable diagnostic insights, which is pivotal for enhancing the accuracy and operational efficiency within the diagnostic landscape. In achieving this, the GUI stands as a testament to our commitment to advancing medical diagnostic capabilities through the synthesis of technology and user-centered design.

External validation

Our recent endeavor extended the clinical validation of our deep learning vascular localization model to incorporate the use of a portable pocket wireless handheld ultrasound device. The external validation was conducted with meticulous adherence to an established compression maneuver protocol, ensuring the reliability and clinical relevance of the results. The study included wireless handheld ultrasound images from 66 patients, resulting in a total of 132 videos used for external validation. Figure 7A illustrates the model's proficiency in accurately predicting the locations of arterial and venous structures, showcasing its robust performance across different imaging environments. This is particularly significant in the context of handheld ultrasound technology, which, while offering portability and convenience, often presents challenges in image quality and consistency.

The validation process employed the SSIM, a metric designed to assess the perceptual quality of visual images. In Figure 7B, an SSIM value in the proximity of 0.9 within our external dataset confirms the model's exceptional accuracy in image prediction. Such high SSIM scores illustrate the model's capability to replicate the precision of traditional, more cumbersome imaging systems within the compact and user-friendly format of handheld devices, indicating a substantial advance in point-of-care imaging diagnostics.

The successful application of our model with portable ultrasound technology marks a significant step forward in expanding diagnostic services to remote and resource-limited settings. By enabling high-fidelity vascular imaging in a diverse array of clinical environments, our model has the potential to significantly improve the efficiency and accessibility of medical diagnoses. This breakthrough holds particular promise for enhancing patient care in underserved areas, potentially transforming the landscape of global health by facilitating earlier detection and treatment of vascular diseases.

Our study implements a network architecture combining UNet and ResNet frameworks for enhanced vessel localization in Figure 8; see STAR Methods for details.

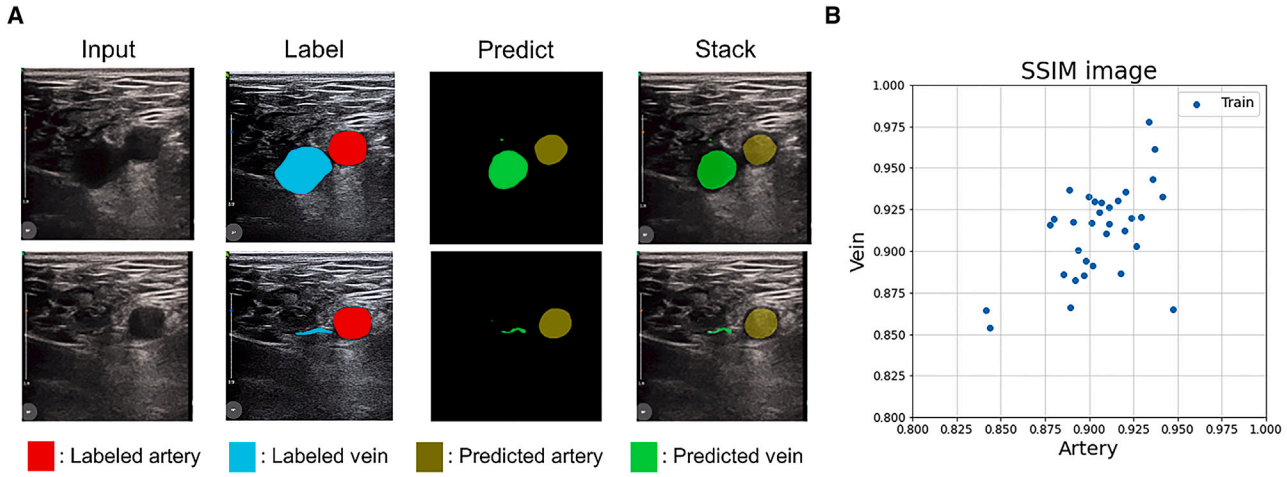


Figure 7. Prediction results from external validation dataset

(A) The deep learning model accurately predicts the images obtained by the portable pocket handheld ultrasound.

(B) Using SSIM to evaluate the similarity between the labeled images and the predicted results, a value of approximately 0.9 is obtained, even with external validation.

DISCUSSION

The incorporation of machine learning into the field of medical imaging has substantially advanced the automated segmentation and recognition of vessel walls and their compressibility, which are vital for the comprehensive management of venous diseases. Our research has validated the efficacy of ResUNet, a specialized machine learning model, which demonstrates remarkable precision in segmenting and discerning vessel wall structures and compressibility from venous duplex ultrasonographic videos. This precision is maintained across imaging modalities obtained via both traditional and portable handheld ultrasound devices, showcasing the adaptability of ResUNet in varying clinical settings.

Machine learning's application extends to the domain of deep vein thrombosis (DVT). A notable study by Jin et al.[24] describes the creation and validation of machine learning models to forecast the risk of cancer-associated DVT. These models were trained using a retrospective analysis of electronic medical records, incorporating a multitude of clinical variables, such as patient baseline characteristics and cancer-associated factors. Significantly, these models were constructed without the incorporation of imaging data, suggesting a potential area of enhancement where image-based machine learning algorithms could contribute to a more holistic risk assessment approach.

Furthermore, the work of Liu et al.[28] explores the utility of a two-dimensional convolutional neural network (2D CNN), specifically the VGG-19 model as proposed by the Visual Geometry Group, in the context of light reflection rheography (LRR). LRR serves as a rapid and non-invasive examination technique that simulates venous pressure hemodynamics, yet it is not widely adopted in standard venous disease protocols. The integration of 2D CNN with LRR demonstrates the burgeoning role of advanced computational models in enhancing the diagnostic accuracy of venous function assessments, reinforcing the potential of machine learning to augment even less common diagnostic protocols.

In the landscape of venous disease diagnostics, anatomical imaging modalities such as venography and venous duplex ultrasonography are deemed more clinically relevant than functional examinations like light reflection rheography (LRR). The differentiation between these modalities lies in their capacity to visualize anatomical details critical for the diagnosis of conditions such as DVT. Contemporary research has leveraged various deep learning models predicated on convolutional neural networks (CNNs) — including VGG16, VGG19, ResNet50, and ResNet152 — to enhance the automated classification of DVT from contrast-enhanced lower extremity computed tomography (CT) and venography. Among these models, VGG16 has emerged as the most performant, as evidenced by its higher Area Under the Curve (AUC) values, while CNN-based ResNet models displayed comparatively lower efficacy.[29] Notably, CT venography is less susceptible to operator-induced variability than venous duplex, although it faces limitations due to the often diminutive size of the blood vessel area, which can be challenging for models like ResNet that require larger initial image sizes for optimal performance.

Our approach differs from previous studies because we based our model on U-Net and integrated it with a ResNet architecture. While other studies proposed a modified U-Net model called Weighted ResUNet with attention mechanisms and ResNet skip connections, our model took a different approach.[26,30] We chose not to include attention mechanisms. After comparing ResNet18, ResNet50, and ResNet101, we found minimal differences in prediction outcomes and ultimately opted for ResNet34. This choice was made to ensure a balance between predictive performance and computational efficiency.

In the domain of image segmentation, UNet has been recognized for its exemplary performance, which is heavily influenced by the number of trainable parameters and its architectural

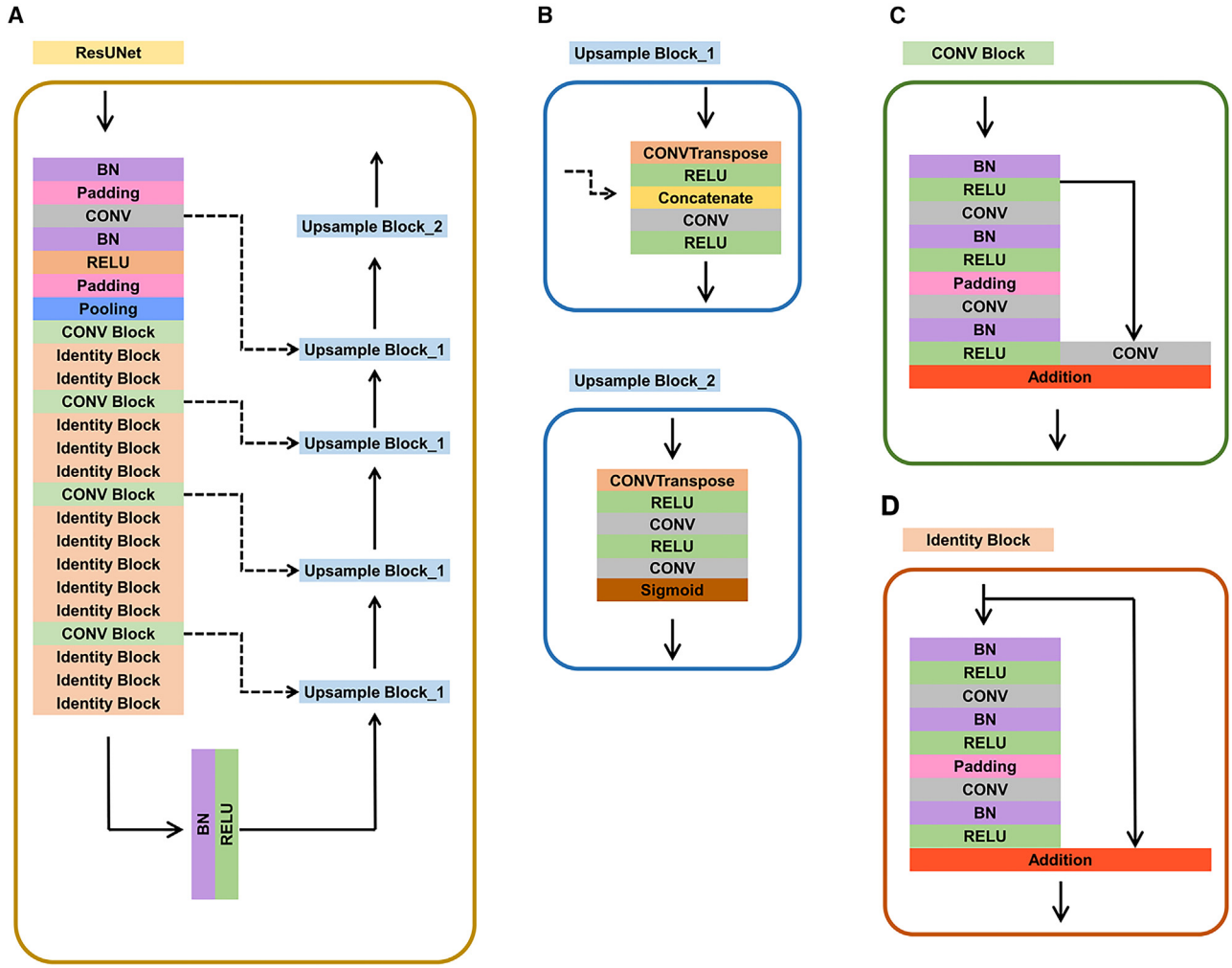


Figure 8. Architectural of the ResUNet model

(A) Architecture of the model.

(B) In the expansive path of UNet, the upsample block concatenates two conv layers with a skip connection and then uses ConvTranspose to complete the upsampling process, which reduced the feature into a new image of the same size as the original image and concatenates to preserve the feature.

(C) The residual blocks include the Conv block to allow the model to learn new features. The use of residual blocks addresses the problem of gradient disappearance in deep learning and improves accuracy.

(D) The identity block maps the features of the previous level directly when the model does not learn new features so that the model will not degrade.

foundation. Nevertheless, its precision is compromised when faced with complex feature geometries and significant background noise.[31] Studies have outlined methods for identifying venous disease using metrics derived from traditional compression ultrasound imagery based on the UNet framework, but these methods suffer from limitations in reproducibility and variance attributable to the ultrasound device manufacturers.[26,27] Conversely, our research indicates that ResUNet, an enhanced version of UNet incorporating residual learning strategies from ResNet, exhibits improved segmentation capabilities, yielding satisfactory results even with the less optimal image quality of portable pocket handheld ultrasound devices.

Lower limb venous ultrasound generally requires relatively lower image quality and technical expertise. Our model achieved the anticipated performance targets without the need for additional denoising techniques. Furthermore, the successful external validation with handheld ultrasound devices confirms that our approach meets the required performance benchmarks even without denoising, aligning with our model's intended use and evaluation criteria. Denoising is indeed an important consideration, as highlighted in previous systematic comparison analysis.[32] As we continue to develop and refine our model for more advanced practical applications, we will explore the integration of pre-processing and denoising techniques to further enhance performance.

Image quality is a pivotal factor in the performance of machine learning models, with higher quality images typically facilitating more precise predictions. Advanced models may necessitate

more complex or deeper network architectures to fully exploit the high-quality data. While UNet is known for its minimal data requirements and tolerance for lower image quality, the volume and quality of data remain crucial for effective model training. The integration of UNet with ResNet, as seen in our ResUNet model, optimizes the learning process and enhances prediction outcomes, effectively addressing challenges associated with image quality during training. Our findings suggest that such hybrid models can be instrumental in improving diagnostic accuracy, particularly in settings where access to high-end imaging equipment is limited.

Conclusion

Our research successfully demonstrated the utility of the ResUNet machine learning model for the automated segmentation and recognition of vessel walls and their compressibility in venous duplex ultrasonographic videos. This was achieved using both standard and portable handheld ultrasound devices. The ResUNet model notably enhanced the performance of the UNet framework in image segmentation, particularly in scenarios involving complex anatomical features and substantial background noise. Impressively, the model yielded satisfactory results even when applied to images of lesser quality, such as those obtained from portable handheld ultrasound devices. This development has significant implications for patient care and the economics of healthcare, offering the possibility for real-time analysis and utilization directly on mobile devices.

Despite the encouraging outcomes, the retrospective design of our study introduces limitations that must be considered. To establish the robustness and general applicability of our findings, further validation is necessary, ideally in the form of a prospective multicenter study. Such studies would benefit from the synergy of industry-academia collaborations, promoting the refinement and practical application of our model across a variety of clinical environments. As machine learning continues to evolve within the medical field, ongoing research is required to further explore and substantiate the viability and dependability of these advanced models for the automated interpretation of medical imaging.

Limitations of the study

While our study has yielded encouraging outcomes, it is imperative to acknowledge certain limitations. The retrospective design may intrinsically carry biases that could influence the results. Therefore, to reinforce the validity of our findings, it would be prudent to pursue a prospective multicenter study, which would enhance the robustness of the model by incorporating a more diverse and extensive dataset. Additionally, external validation with datasets from various institutions would be beneficial to ascertain the model's reliability and generalizability.

Despite these limitations, the echocardiographic equipment utilized in our study is representative of the machines commonly employed in clinical settings, suggesting that our results have a high potential for applicability in standard practice. Furthermore, the standard echocardiographic views that our model relies upon are uniformly adopted in routine examinations, and our model's performance has been corroborated by an external dataset obtained from a portable handheld ultrasound device.

To our knowledge, this study is pioneering in demonstrating the feasibility of applying a machine-learning model to portable medical imaging devices. This advancement holds considerable promise for real-time diagnostic analysis and the potential for integration into mobile device platforms, which could significantly streamline patient care and optimize healthcare expenditure.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, C.-H.Y. (jonnyyu@gs.ncku.edu.tw).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- The complete original data reported in this study cannot be deposited in a public repository because these data are confidential medical records. The dataset used and analyzed during the current study is available upon reasonable request, contact C.-H.Y. (jonnyyu@gs.ncku.edu.tw).
- The original methodologies and code implementations of UNet and ResNet used in this study were referenced from the original papers, which are listed in the key resources table. Model pretrained weights for ResNet34 have been deposited at GitHub (https://github.com/qubvel/classification_models). This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

This study was financially supported by the National Science and Technology Council [NSTC, formerly the Ministry of Science and Technology (MOST) in Taiwan MOST 111-2314-B-006-047]. This study was funded by research grants from the National Cheng Kung University Hospital in Taiwan (NCKUH-11203019, NCKUH-11303021). The authors would like to express their gratitude to the NCKUH. P.W.C. and P.Y.L. were also financially supported by the NSTC, formerly MOST, in Taiwan with grant numbers 111-2314-B-006-017-MY3, 108-2314-B-006-098-MY3, and 111-2634-F-006-007 and the grant of D111-G2512 from the Higher Education Sprout Project, Ministry of Education to the Headquarters of University Advancement at National Cheng Kung University. We would like to thank Anthony Abram (www.uni-edit.net) for editing and proofreading this manuscript. We also acknowledged the contribution by Distinguished Professor Huey-Jen Jenny Su and Mong-Run Shen from National Cheng Kung University for their leaderships and grant aids from MOST AI Capstone program for the establishing pipeline of analysis for clinical images.

AUTHOR CONTRIBUTIONS

Conceptualization: P.W.C., C.H.Y., and P.Y.L.; Data curation: P.W.C., B.Y.T., and Z.H.Y.; Formal analysis: P.W.C., B.Y.T., Z.H.Y., K.T.L., and J.N.C.; Funding acquisition: P.W.C., C.H.Y., and P.Y.L.; Investigation: C.H.Y. and P.Y.L.; Methodology: P.W.C., B.Y.T., Z.H.Y., K.T.L., J.N.C., and C.H.Y.; Project administration: P.W.C., B.Y.T., Z.H.Y., and C.H.Y.; Software: P.W.C., B.Y.T., Z.H.Y., and C.H.Y..; Supervision: C.H.Y. and P.Y.L.; Writing – original draft: P.W.C., B.Y.T., and Z.H.Y.; Writing – review and editing: P.W.C., C.H.Y., and P.Y.L.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Ethics
- METHOD DETAILS
  - Venous duplex protocol
  - Dataset
  - Network architecture
  - Training process
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.111318.

Received: February 4, 2024
Revised: July 14, 2024
Accepted: October 30, 2024
Published: November 4, 2024

REFERENCES

1. Kakkos, S.K., Gohel, M., Baekgaard, N., Bauersachs, R., Bellmunt-Montoya, S., Black, S.A., Ten Cate-Hoek, A.J., Elalamy, I., Enzmann, F.K., Geroulakos, G., et al. (2021). Editor's choice–European Society for Vascular Surgery (ESVS) 2021 clinical practice guidelines on the management of venous thrombosis. Eur. J. Vasc. Endovasc. Surg. *61*, 9–82.
2. Wang, K.-L., Chu, P.-H., Lee, C.-H., Pai, P.-Y., Lin, P.-Y., Shyu, K.-G., Chang, W.-T., Chiu, K.-M., Huang, C.-L., Lee, C.-Y., et al. (2016). Management of venous thromboembolisms: part I. The consensus for deep vein thrombosis. Acta Cardiol. Sin. *32*, 1–22.
3. Wang, K.-L., and Chu, P.-H. (2021). Management of Venous Thromboembolisms: Part II. The Consensus for Pulmonary Embolism and Updates. Acta Cardiol. Sin. *37*, 215–216.
4. Di Nisio, M., van Es, N., and Büller, H.R. (2016). Deep vein thrombosis and pulmonary embolism. Lancet *388*, 3060–3073.
5. Pedraza García, J., Valle Alonso, J., Ceballos García, P., Rico Rodríguez, F., Aguayo López, M.Á., and Muñoz-Villanueva, M.D.C. (2018). Comparison of the accuracy of emergency department-performed point-of-care-ultrasound (POCUS) in the diagnosis of lower-extremity deep vein thrombosis. J. Emerg. Med. *54*, 656–664.
6. Caprini, J.A., Arcelus, J.I., and Reyna, J. (2001). Effective risk stratification of surgical and nonsurgical patients for venous thromboembolic disease. In Seminars in hematology (Elsevier), pp. 12–19.
7. Lim, W., Le Gal, G., Bates, S.M., Righini, M., Haramati, L.B., Lang, E., Kline, J.A., Chasteen, S., Snyder, M., Patel, P., et al. (2018). American Society of Hematology 2018 guidelines for management of venous thromboembolism: diagnosis of venous thromboembolism. Blood Adv. *2*, 3226–3256.
8. Needleman, L., Cronan, J.J., Lilly, M.P., Merli, G.J., Adhikari, S., Hertzberg, B.S., DeJong, M.R., Streiff, M.B., and Meissner, M.H. (2018). Ultrasound for lower extremity deep venous thrombosis: multidisciplinary recommendations from the Society of Radiologists in Ultrasound Consensus Conference. Circulation *137*, 1505–1515.
9. Adhikari, S., Zeger, W., Thom, C., and Fields, J.M. (2015). Isolated deep venous thrombosis: implications for 2-point compression ultrasonography of the lower extremity. Ann. Emerg. Med. *66*, 262–266.
10. Clark, R.M., Weingardt, D., Goff, J.M., Jr., and Ketteler, E.R. (2019). Effects of a standardized emergency department protocol on after-hours use of venous duplex ultrasound. J. Vasc. Surg. Venous Lymphat. Disord. *7*, 501–506.
11. Deagle, J., Allen, J., and Mani, R. (2005). A nurse-led ambulatory care pathway for patients with deep venous thrombosis in an acute teaching hospital. Int. J. Low. Extrem. Wounds *4*, 93–96.
12. Johnson, K.W., Torres Soto, J., Glicksberg, B.S., Shameer, K., Miotto, R., Ali, M., Ashley, E., and Dudley, J.T. (2018). Artificial intelligence in cardiology. J. Am. Coll. Cardiol. *71*, 2668–2679.
13. Zeng, Z., Wang, Q., Yu, Y., Zhang, Y., Chen, Q., Lou, W., Wang, Y., Yan, L., Cheng, Z., Xu, L., et al. (2022). Assessing electrocardiogram changes after ischemic stroke with artificial intelligence. PLoS One *17*, e0279706.
14. Manlhiot, C., Van den Eynde, J., Kutty, S., and Ross, H.J. (2021). A primer on the present state and future prospects for machine learning and artificial intelligence applications in cardiology. Can. J. Cardiol. *38*, 169–184.
15. Yu, C., Tseng, B., Yang, Z., Tung, C., Zhao, E., Ren, Z., Yu, S., Chen, P., Chen, C., and Buehler, M.J. (2022). Hierarchical Multiresolution Design of Bioinspired Structural Composites Using Progressive Reinforcement Learning. Adv. Theory Simul. *5*, 2200459.
16. Yu, C.-H., Wu, C.-Y., and Buehler, M.J. (2022). Deep learning based design of porous graphene for enhanced mechanical resilience. Comput. Mater. Sci. *206*, 111270.
17. Hsu, Y.-C., Yu, C.-H., and Buehler, M.J. (2021). Tuning mechanical properties in polycrystalline solids using a deep generative framework. Adv. Eng. Mater. *23*, 2001339.
18. Althnian, A., AlSaeed, D., Al-Baity, H., Samha, A., Dris, A.B., Alzakari, N., Abou Elwafa, A., and Kurdi, H. (2021). Impact of dataset size on classification performance: an empirical evaluation in the medical domain. Appl. Sci. *11*, 796.
19. Saeedizadeh, N., Minaee, S., Kafieh, R., Yazdani, S., and Sonka, M. (2021). COVID TV-Unet: Segmenting COVID-19 chest CT images using connectivity imposed Unet. Comput. Methods Programs Biomed. Update *1*, 100007.
20. Guan, S., Khan, A.A., Sikdar, S., and Chitnis, P.V. (2020). Fully Dense UNet for 2-D Sparse Photoacoustic Tomography Artifact Removal. IEEE J. Biomed. Health Inform. *24*, 568–576.
21. Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (Springer), pp. 234–241.
22. He, K., and Sun, J. (2015). Convolutional neural networks at constrained time cost. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 5353–5360.
23. He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778.
24. Jin, S., Qin, D., Liang, B.-S., Zhang, L.-C., Wei, X.-X., Wang, Y.-J., Zhuang, B., Zhang, T., Yang, Z.-P., Cao, Y.-W., et al. (2022). Machine learning predicts cancer-associated deep vein thrombosis using clinically available variables. Int. J. Med. Inform. *161*, 104733.
25. Fong-Mata, M.B., García-Guerrero, E.E., Mejía-Medina, D.A., López-Bonilla, O.R., Villarreal-Gómez, L.J., Zamora-Arellano, F., López-Mancilla, D., and Inzunza-González, E. (2020). An artificial neural network approach and a data augmentation algorithm to systematize the diagnosis of deep-vein thrombosis by using wells' criteria. Electronics (Basel) *9*, 1810.
26. Kainz, B., Heinrich, M.P., Makropoulos, A., Oppenheimer, J., Mandegaran, R., Sankar, S., Deane, C., Mischkewitz, S., Al-Noor, F., Rawdin, A.C., et al. (2021). Non-invasive diagnosis of deep vein thrombosis from ultrasound imaging with machine learning. NPJ Digit. Med. *4*, 137.
27. Tanno, R., Makropoulos, A., Arslan, S., Oktay, O., Mischkewitz, S., Al-Noor, F., Oppenheimer, J., Mandegaran, R., Kainz, B., and Heinrich, M.P. (2018). Autodvt: Joint real-time classification for vein compressibility analysis in deep vein thrombosis ultrasound diagnostics. In International

Conference on Medical Image Computing and Computer-Assisted Intervention (Springer), pp. 905–912.

28. Liu, S.-H., Chen, W., Su, C.-H., and Pan, K.-L. (2022). Convolutional neural Network-based detection of deep vein thrombosis in a low limb with light reflection rheography. Measurement *189*, 110457.
29. Hwang, J.H., Seo, J.W., Kim, J.H., Park, S., Kim, Y.J., and Kim, K.G. (2022). Comparison between Deep Learning and Conventional Machine Learning in Classifying Iliofemoral Deep Venous Thrombosis upon CT Venography. Diagnostics *12*, 274.
30. Xiao, X., Lian, S., Luo, Z., and Li, S. (2018). Weighted res-unet for high-quality retina vessel segmentation. In 2018 9th international conference on information technology in medicine and education (ITME) (IEEE), pp. 327–331.
31. Sharma, P., Ninomiya, T., Omodaka, K., Takahashi, N., Miya, T., Himori, N., Okatani, T., and Nakazawa, T. (2022). A lightweight deep learning model for automatic segmentation and analysis of ophthalmic images. Sci. Rep. *12*, 8508.
32. Ardakani, A.A., Mohammadi, A., Faeghi, F., and Acharya, U.R. (2023). Performance evaluation of 67 denoising filters in ultrasound images: A systematic comparison analysis. Int. J. Imaging Syst. Technol. *33*, 445–464.
33. Guideline developed in collaboration with the American College of Radiology; Society of Pediatric Radiology; Society of Radiologists in Ultrasound (2015). AIUM Practice Guideline for the Performance of Peripheral Venous Ultrasound Examinations. J. Ultrasound Med. *34*, 1–9.
34. Eriksson, F.S., Broms, M.S.A., and Bagge, M.S.J. (2019). Diagnosing Pneumothorax in X-Ray Images Using Deep Learning.
35. Milletari, F., Navab, N., and Ahmadi, S.-A. (2016). V-net: Fully convolutional neural networks for volumetric medical image segmentation. In 2016 fourth international conference on 3D vision (3DV) (IEEE), pp. 565–571.
36. Wang, Z., Bovik, A.C., Sheikh, H.R., and Simoncelli, E.P. (2004). Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. *13*, 600–612.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Python version 3.7	Python Software Foundation	https://www.python.org
TensorFlow version 2.5	TensorFlow Software Foundation	https://www.tensorflow.org
UNet	Ronneberger et al.[21]	https://doi.org/10.1007/978-3-319-24574-4_28
ResNet	He et al.[23]	https://doi.org/10.1109/CVPR.2016.90
Model pretrained weights for ResNet34	Pavel Iakubovskii	https://github.com/qubvel/classification_models

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

We conducted a retrospective study involving patients who underwent venous duplex imaging at our echo lab from January 2021 to August 2022. The data collection and protocols used in this study were approved by the Medical Ethics Committee of National Cheng-Kung University Hospital [IRB: A-ER-111-331]. This study included 152 consecutive patients examined for proximal DVT detection. Data were collected from venous duplex examinations at National Cheng Kung University Hospital in southern Taiwan, used for both training and testing sets. Labels were obtained from examination reports prepared by research assistants and reviewed by cardiologists who authorized the final reports. All images were manually reviewed by two blinded specialists in our peripheral lab. The ultrasound scans were conducted by experienced research assistants and cardiologists, resulting in 570 fully labeled vascular ultrasound images. Clinical data for the training dataset were acquired from our electronic medical records. The baseline characteristics of enrolled patients are detailed in Table S1. The performance of our model remained consistent regardless of participants' sex and gender.

Ethics

The data collection and protocols used in this study were conducted according to the principles in the Declaration of Helsinki and were approved by the Medical Ethics Committee of National Cheng-Kung University Hospital [IRB: A-ER-111-331].

METHOD DETAILS

Venous duplex protocol

All venous duplex examinations were performed in our echo lab, using the ACUSON P500TM ultrasound system (Siemens Medical Solutions USA Inc.) with VF13-5 linear transducers (4.1–12.1 MHz). Transverse transducer compressions were performed for bilateral femoral and popliteal vein. Dynamic images during real-time manual compression were all recorded as dynamic programs. Following our standard imaging algorithm, complete venous duplex evaluation incorporated B mode and spectral Doppler with or without color flow Doppler analysis. When thrombus or intraluminal echoes were present, the location and extent of thrombus with differentiation between partially and totally non-compressible segments were also performed.[8,33]

After de-identification, we constructed a new image database for peripheral venous duplex, involving bilateral femoral/popliteal vein B-Mode images, degree of compressibility, color Doppler images, spectral Doppler waveforms, and any abnormal findings.

For external validation, portable pocket wireless and handheld ultrasounds were performed using the LeSONO LU700L ultrasound system (LELTEK Inc., Taiwan) with a linear transducer (5–10 MHz). Among a subset of the enrolled subjects, portable handheld ultrasounds are concurrently employed in our echo lab.

Dataset

The dataset employed in this study consisted of 570 ultrasound images, segmented into distinct sets for training (428 images), validation (84 images), and testing (58 images). In an effort to enhance the model's performance significantly, a technique involving the horizontal flipping of images was utilized. This method effectively augmented the training dataset, thereby enriching the variability of the dataset without necessitating the labor-intensive process of additional data collection and annotation.

In the preliminary phase of data preprocessing, the initial step involves the strategic cropping of the ultrasound images. The objective of this procedure is to isolate the pertinent segment of each image, concurrently removing extraneous noise which might otherwise introduce undesirable variables into the training process. Following this, the application of LabelMe, an annotation tool, is undertaken to meticulously label the positions of arterial and venous structures. This step ensures the provision of precise and accurate annotation data, which is crucial for the subsequent phases of model training and analytical evaluation.

Network architecture

In this research, we have developed an innovative model that synergistically combines the architectures of UNet and ResNet to achieve precise localization of blood vessels. The UNet architecture is dichotomized into two principal components: the Encoder and the Decoder, as delineated in Figure 8A. The Encoder's primary role is to downsample the image, thereby extracting essential features and simultaneously diminishing noise. In contrast, the Decoder, also referred to as the Expansion Path, is tasked with restoring the image to its original dimensions through a process known as deconvolution. A significant challenge inherent in this method is the potential misidentification of fine image features as noise during the downsampling phase, which can result in their inadvertent loss.

To circumvent this issue, the UNet framework employs a method of concatenated downsampling and upsampling, commonly referred to as 'skip connection', as depicted in Figure 8B. This approach is of critical importance in segmentation tasks, wherein the accuracy of the model's predictions is intimately connected to the spatial arrangement of the features in the input image.

UNet is renowned for its exceptional ability in image segmentation, facilitating the accurate identification of specific targets within an image. Nonetheless, as the depth of the network increases, so does the complexity of the learning process, which may lead to a degradation in accuracy. To address this, our model integrates ResNet with UNet during the downsampling phase. This integration is particularly focused on the accurate prediction of arterial and venous structures in ultrasound images. The architecture comprises a convolution block and two to five identity blocks merged into a singular downsampling block, each incorporating a residual block, as illustrated in Figures 8C and 8D. Overall, the model we have developed demonstrates exceptional proficiency in the accurate prediction of ultrasound images, effectively addressing the challenges associated with fine feature recognition and the potential for accuracy degradation in deeper network architectures.

The ResNet architecture uses the concept of residual learning to solve the problem of accuracy saturation and rapid decline in accuracy when network depth increases. This approach trains the network to learn the difference between the input and the feature, also known as the residual, in a stacked layer structure. The incorporation of residual learning has proven to be highly effective in preventing accuracy saturation and improving performance. For a stacked layer structure, when the input is x, the learned feature is denoted as $H(x)$; and we expect that it can learn the difference between the input and the feature, with the difference here being the residual. The formula is shown as Equation 1:

$$F(x) = H(x) - x. \quad \text{(Equation 1)}$$

The original learned feature, $H(x)$, can be represented as $F(x) + x$. Residual learning is preferred over using the original feature directly. When the residual is zero, the accumulation layer only performs identity mapping. This ensures that, at minimum, the network performance does not decrease. However, it is rare for the residual to be exactly zero as it generally contains information that can lead to improved performance in the accumulation layer. This makes residual learning highly effective at achieving high accuracy and improving network performance.

Training process

In this study, we partitioned the dataset into distinct sets for training, validation, and testing, constituting 75%, 15%, and 10% of the total data, respectively. The training of the model was conducted on a GeForce RTX 3090Ti GPU, utilizing a batch size of 32 over the course of 5000 epochs. For the optimization process, the Adam optimizer was deployed, alongside a hybrid loss function integrating Binary Cross-Entropy with Dice Loss. This combination facilitated the model's enhanced capability in discerning both macroscopic and microscopic anatomical structures within ultrasound images. The meticulous design of our training and testing methodologies has been instrumental in the establishment of a model characterized by high precision and efficiency. This model is particularly adept at delineating the locations of blood vessels in ultrasonographic scans, thereby demonstrating its potential utility in medical imaging diagnostics.

QUANTIFICATION AND STATISTICAL ANALYSIS

The loss function is used to evaluate the degree of inconsistency between the predictions and the ground truth of the model: a lower loss function means that the model's prediction is more accurate. We used binary cross entropy and log dice loss in combination as a loss function. Binary cross-entropy is used to evaluate how well or how poorly a binary category model predicts results and is defined as Equation 2:

$$L_{BCE}(x_i,\tilde{x}_i) = -\frac{1}{N}\sum_{i=1}^{N} x_i \log(\tilde{x}_i) + (1 - x_i)\log(1 - \tilde{x}_i) \quad \text{(Equation 2)}$$

where x_i is the ground truth pixel value and $\tilde{x}_i$ is the predicted pixel values. In the case that $x_i = 1$, if the predicted value of $\tilde{x}_i$ tends to one, then the value of the loss function should tend to zero. Conversely, if the predicted value $\tilde{x}_i$ tends to zero at this point, then the value of the loss function should be very high. The log dice loss is defined as Equation 3:

$$L_{LogDice}(x_i, \tilde{x}_i) = -\log\left(\frac{2\sum_{i=1}^{N} x_i\tilde{x}_i + Smooth}{\sum_{i=1}^{N} x_i + \sum_{i=1}^{N} \tilde{x}_i + Smooth}\right) \quad \text{(Equation 3)}$$

Log-Dice is a loss function based on the Dice coefficient, which is used to observe the overlapping areas of the labeled and predicted images. The value will be closer to one when there are more overlapping areas.[34]

We also used the dice coefficient as a quantitative indicator for model evaluation, which is defined as Equation 4:

$$Dice = \frac{2 \times TP}{(TP + FP) + (TP + FN)} \quad \text{(Equation 4)}$$

True Positive (TP) represents when the model predicts an artery at its image position; False Negative (FN) represents when the model does not predict an artery at its image position; False Positive (FP) represents when the model predicts an artery at an image position at which there is no artery.[35]

Finally, we also used the Structural Similarity Index (SSIM) to evaluate the similarity between the labeled images and the predicted images. SSIM is mainly used for comparing the brightness, contrast, and structure between two images, and is defined as Equation 5:

$$SSIM(x, y) = \frac{(\mu_x\mu_y + c_1)(\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad \text{(Equation 5)}$$

where μ_x and μ_y represent the average of x and y respectively; σ_x and σ_y represent the standard deviation of x and y respectively; σ_{xy} represents the covariance of x and y; and c_1 and c_2 are constants. SSIM is between 0 and 1. The higher the SSIM, the smaller the difference between the two images. When the two images are the same, the SSIM will be equal to one.[36]