**The Version of Record of this manuscript has been published and is available in *Quantitative InfraRed Thermography Journal* (2025): https://www.tandfonline.com/doi/full/10.1080/17686733.2025.2585417**

# Optimizing Neural Networks for Perforator Detection in DIEP Flap Breast Reconstruction Using Dynamic Infrared Thermography

by Warre Clarys*, Rhys Evans*, Simon Verspeek*, Jan Verstockt*, Hai Zhang*****,Veronique Verhoeven****, Wiebren A.A. Tjalma***, Filip Thiessen**, Gunther Steenackers*

* *InViLab Research Group, University of Antwerp, Groenenborgerlaan 171, B-2020 Antwerp, Belgium*

** *Department of Plastic, Reconstructive and Aesthetic Surgery, Multidisciplinary Breast Clinic, Antwerp University Hospital, University of Antwerp, Wilrijkstraat 10, B-2650 Antwerp, Belgium*

****Gynaecological Oncology Unit, Department of Obstetrics and Gynaecology, Multidisciplinary Breast Clinic, Antwerp University Hospital, Wilrijkstraat 10, B-2650 Antwerp, Belgium*

*****Department of Family Medicine and Population Health, University of Antwerp, Antwerp, Belgium*

******Department of Electrical and Computer Engineering, Computer Vision and Systems Laboratory (CVSL), Laval University, Quebec, G1V 0A6, Canada.*

*Correspondance: warre.clarys@uantwerpen.be*

**Abstract**

Breast reconstruction following mastectomy is increasingly performed, with Deep Inferior Epigastric artery Perforator (DIEP) flap surgery considered the gold standard. Accurate preoperative perforator selection is vital to minimize complications and operative time. While computed tomography angiography (CTA) remains the clinical reference, drawbacks including radiation, contrast use, and cost motivate exploration of non-invasive alternatives. Dynamic Infrared Thermography (DIRT) offers a low-cost, radiation-free method but still lacks automation.

This study evaluates deep learning for automated perforator detection in DIRT. A dataset of 50 time-lapse thermograms from five patients was acquired using various cooling methods and validated through leave-one-out cross-validation (LOOCV). Two neural network architectures were compared: a standard U-Net and a modified U-Net (mU-Net) from prior work.

U-Net consistently outperformed mU-Net. Across LOOCV folds, U-Net achieved a mean weighted Dice loss of 0.42 ± 0.09, sensitivity of 0.87 ± 0.08, and precision of 0.82 ± 0.14. On an independent test patient, sensitivity remained high (0.88) but precision decreased (0.58). The mU-Net failed to converge (validation loss 0.87 ± 0.02), producing uniform segmentations.

These findings demonstrate that U-Net is a robust tool for automated perforator detection in DIRT, though false positives highlight the need for larger datasets and further optimization before clinical use.

**Keywords:** DIEP flap surgery, dynamic infrared thermography, perforator detection, neural network, U-Net architecture

## 1. Introduction

In 2020, breast cancer became the most common cancer worldwide, with an estimated 2.3 million new cases. About one in seven women in Western countries will encounter breast cancer during their lifetime [1]. Moreover, predictions indicate that breast cancer will experience the most substantial increase among all cancers. According to the International Agency for Research on Cancer of the World Health Organization, there is an expected 92.8% rise in breast cancer cases by 2045 [2].

Over recent decades, there has also been a notable increase in breast reconstructions, particularly in Western regions. Studies conducted between 2005 and 2014 revealed that approximately 35-40% of breast cancer cases in the USA undergo mastectomy, with 50.5% of them opting for breast reconstruction [3], [4]. This accounts for about 18% of post-mastectomy breast reconstructions relative to breast cancer cases. In recent years, this proportion has significantly risen, reaching 55.3% in 2022 [2], [5]. This remarkable surge in breast reconstructions can be attributed to various factors, including improved surgical techniques, greater acceptance of reconstructive options, and a growing number of prophylactic mastectomies [6]. Additionally, studies confirm the positive impact of breast reconstruction on the psychological well-being and quality of life post-mastectomy [6], [7].

Regarding breast reconstruction methods, there is a growing preference for autologous reconstructions using the patient's own tissue. Autologous reconstructions offer advantages such as a more natural appearance and feel of the breast, with longer-lasting results compared to implants. Deep Inferior Epigastric artery Perforator (DIEP) flap surgery has emerged as the gold standard for autologous breast reconstructions [8], [9]. During a DIEP flap breast reconstruction procedure, skin and subcutaneous tissue from the lower abdomen of the patient are utilized to reconstruct the breast. This tissue, known as a flap, receives its blood supply from the deep inferior epigastric artery and one or two associated veins through a perforator. During the surgery, the perforator is anastomosed to the internal mammary artery and vein to ensure optimal blood flow to the reconstructed breast [10]. Selecting the most

suitable perforators is crucial for optimizing the effectiveness of this procedure, leading to reduced necrosis and complication rates, improved outcomes and shorter operation times [10], [11].

Currently, several techniques are used to locate perforator vessels, such as computed tomography angiography (CTA), color Doppler ultrasonography (CDU), magnetic resonance angiography (MRA), or dynamic infrared thermography (DIRT) [11]. Among these, CTA is considered the gold standard for perforator selection, as it provides detailed information about the location and intramuscular course of the perforators. However, CTA has multiple disadvantages, such as the use of intravenous contrast agents and ionizing radiation, which can damage genetic material and potentially lead to cancer and kidney failure [12], [13]. Other drawbacks are the high costs associated with CTA scans, as well as being unsuitable for intraoperative use and lacking information about the physiological information in flow characteristics of perforators pre- and postoperatively [10].

An alternative, non-invasive technique for perforator selection is Dynamic Infrared Thermography (DIRT). Unlike CTA scans, DIRT does not require the use of ionizing radiation or contrast agents. Although initially introduced to reconstructive surgeries by Theuvenet et al. in 1986, this technique has gained popularity in recent years [14]. It is an imaging technique that measures the temperature changes in the body's surface tissues by detecting infrared (IR) radiation after an external stimulus. DIRT is a cost-effective method, as the major expense lies in the thermal camera, which is ten to one hundred times cheaper than a CT scanner. This technique is utilized in diverse medical applications, including breast reconstructions, breast cancer, skin cancer, and diabetes screening [15]. Typically, the skin is cooled, after which the tissue self-reheats due to blood circulation [16], [17]. The IR measurements taken during the rewarming process provide the desired information about the tissue, depending on the application. DIRT is used preoperatively in DIEP flap surgeries to optimize the selection of perforators by identifying dominant perforators through analysis of hotspots during the rewarming phase [18], [19].

Previous efforts to automate perforator selection in DIEP flap surgery have primarily aimed at reducing labor intensity and interpretation time. Mavioso et al. [20] introduced a method based on CTA and conventional segmentation techniques, demonstrating the potential to improve accessibility and procedural efficiency by reducing reliance on radiologists. Similarly, Kakileti et al. [21] focused on automating breast cancer screening through thermography, underscoring the importance of standardized quantification to mitigate subjective interpretation. Despite their contributions, both approaches depended on segmentation techniques sensitive to noise, lighting, and external factors. While intensity and area thresholding methods may perform well under controlled conditions, they lack robustness in real-world applications, often yielding errors when confronted with variations in temperature, noise, or camera angles. Moreover, abdominal temperature varies among patients and can be influenced by ambient room temperature, further complicating accurate segmentation. Recognizing these limitations, Mavioso proposed that machine learning techniques could enhance robustness. Building on this premise, De La Hoz et al. [22] recently developed a neural network model for perforator detection, achieving high segmentation accuracy with a Dice score of 0.8 after 2 minutes of rewarming and 0.9 after 4 minutes. However, the current validation was performed using a limited dataset of 27 measurements. De La Hoz et al. [22] concluded that training and validation on larger, more diverse datasets could further improve the model's generalizability and robustness.

This study aims to evaluate the accuracy and robustness of the neural network model developed by De La Hoz et al. [22] using a larger and more diverse dataset. Additionally, it is of particular interest to assess perforator detection using a U-Net neural network, which is widely employed in the medical domain for segmentation tasks. U-Net has proven effective in applications such as liver lesion segmentation [23] and retinal vessel detection [24], demonstrating its adaptability to complex medical imaging challenges. In this paper, the performance of the two neural networks will be compared to determine their efficacy in perforator detection. The overarching goal of this research is to highlight the importance of representative datasets, advanced preprocessing techniques, and optimized neural network architectures in enhancing abdominal perforator detection accuracy in DIEP flap surgeries.

This study is organized into six sections. Section 2 outlines the methodology, detailing the data acquisition process, the neural network architectures, and the training protocols employed. Section 3 focuses on the performance analysis, describing the evaluation metrics and the leave-one-out cross-validation approach used to assess the models. In Section 4, the results are presented, with a comparison of the U-Net and mU-Net architectures in terms of segmentation accuracy and robustness. Section 5 provides a critical discussion of the findings, including their implications, limitations, and potential directions for future research. Finally, Section 6 summarizes the key conclusions and contributions of this work.

## 2. Methodology

### *2.1 Measurements*

In this study, we adhere to the dynamic thermography protocol outlined by Thiessen et al. [25]. Ten different preoperative Dynamic Infrared Thermography (DIRT) measurements were performed per patient across a cohort of five individuals scheduled for DIEP flap surgeries. A FLIR A655sc thermal camera with a 25° lens was employed for data collection. The camera features an uncooled, long-wavelength microbolometer detector (7.5–14 μm), a

spatial resolution of 640 × 480 pixels, and high thermal sensitivity (NETD < 30 mK), enabling precise thermal imaging for clinical applications. The camera was positioned at a distance of 0.9 to 1.2 meters above the patient's lower abdomen, with the exact positioning adjusted according to the surgical table height and the patient's body size.

For each patient, the lower abdomen was subjected to ten distinct cooling techniques to establish a robust and diverse dataset, thereby reducing reliance on a single cooling method. These techniques included conductive and convective approaches, each designed to optimize cooling effectiveness and minimize artifacts such as uncooled creases, which can compromise rewarming measurements. Conductive methods involved direct contact with the skin and included standard water-based cooling using a sterile water-filled bag, as well as variations with insulation materials such as cotton, neoprene, or thermal pads to improve temperature distribution. Elastic cooling was achieved using a water-filled latex balloon, while cryogenic air-filled bags, pressed gently against the abdomen, enhanced cooling uniformity by venting air to prevent localized overcooling. Pre-cooled pads, set to specific temperatures (5 °C or -10 °C), provided additional controlled conductive cooling options. Convective cooling methods used cryogenic air streams directed onto the abdomen at a fixed distance, either stationary or moving, ensuring even cooling without direct skin contact. Skin rewarming was recorded over a 4-minute period after each cooling application to evaluate the thermal effects of each technique.

In total, a dataset comprising 50 time-lapse thermograms was generated. By employing multiple cooling techniques, this study seeks to create a more generalized and robust framework for perforator detection, mitigating the dependence on specific cooling strategies typically chosen by surgeons.

### *2.2 Labeling and augmenting*

The labeling of the dataset was performed using standard threshold-based image segmentation on three images captured during the rewarming phase of the lower abdomen: immediately after cooling (1), after 120 seconds (2), and after 240 seconds (3) (see Figure 1). This approach leverages the advantage of the temporal dimension: immediately after cooling, creases are most visible, while perforators become progressively clearer during rewarming. After 120 seconds, the first perforators are visible, and by 240 seconds, all perforators are fully distinguishable [26]. The thresholding tool was validated using CTA scans of the patients and reviewed by clinical experts to ensure accuracy. The dataset was segmented into four distinct classes, following a method similar to that described by De La Hoz et al. [22]. These four classes (cooled abdomen, umbilicus, perforators, and cooling artifacts) form the basis for both the ground truth labels and the output predictions of the neural networks (see Figure 2). Class 1 represents the cooled abdomen, which serves as the primary region within which all other classes are located. This region is segmented by identifying low-intensity areas immediately after cooling. Class 2 is the umbilicus, a stable, high-temperature region at the upper center of the cooled abdomen, acting as a critical reference point for localization and comparison with CTA scans. Class 3 represents perforators, which are high-intensity regions that emerge during rewarming and exhibit initial cooling. Perforators that are closest to the umbilicus are prioritized, as they are considered optimal for transplantation. Finally, Class 4 includes cooling artifacts, which are uneven cooling lines that initially appear brighter than the surrounding abdomen but gradually converge with the surrounding tissue over time. These were separately labeled to minimize false positives in perforator detection.

Compared to prior work, this research fully utilized the advantage of the temporal dimension, derived from the recorded data, by using the first frame once cooling has been removed, a frame in the middle of the recording and the last frame of the recording. These three frames give the neural network more dimensional context to improve its performance and to make the model more robust.

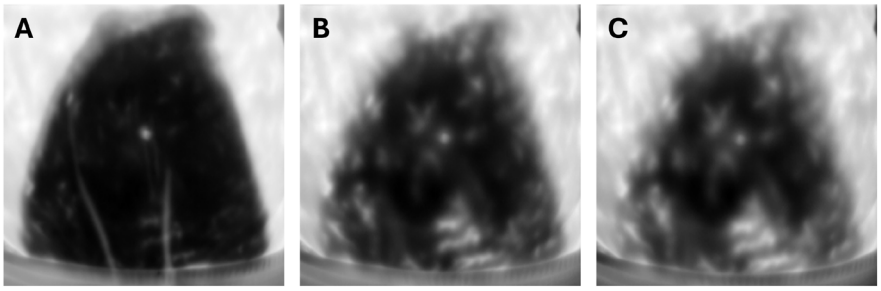



*Figure 1. Three images from the rewarming phase of the lower abdomen are used for labeling: immediately after cooling (A), after 120 seconds (B), and after 240 seconds (C).*

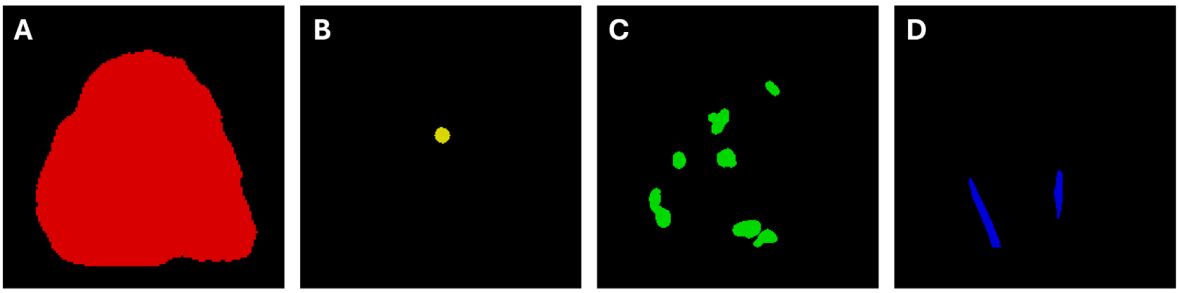



*Figure 2. The dataset is segmented into four distinct classes: cooled abdomen (A), umbilicus (B), perforators (C), and cooling artifacts (D).*

Data augmentation was applied on the 50 measurements to increase the diversity of the training dataset and improve the robustness of the model. The augmentation process used random cropping with a minimum coverage of 70%, resizing all cropped images to a fixed resolution of 256x256, added random Gaussian noise with a standard deviation of 0.25, and performed random horizontal and vertical flips. Multiple augmented copies of each original image were generated by repeating these transformations, thereby significantly enhancing the variability of the dataset. This precise approach reduces the risk of overfitting and improves the model's ability to generalize to unseen data.

### *2.3 Training and validating*

To evaluate the performance and robustness of the neural networks, a comprehensive leave-one-out cross-validation (LOOCV) approach was applied at the patient level. In this method, each patient's dataset is used as the validation set once, while the remaining datasets are used for training. This ensures that each patient is independently validated, minimizing overfitting to specific anatomical characteristics and providing a reliable measure of generalizability. Performance was assessed by comparing training and validation loss across the five cross-validation iterations: a lower validation loss reflects better generalization to unseen data, while a small gap between training and validation loss indicates that the network avoids overfitting. By systematically validating each patient, this approach not only identifies the better-performing network architecture but also evaluates its consistency across different patient anatomies and cooling conditions. In addition to the LOOCV evaluation, a sixth patient was reserved as an independent external test case. For this evaluation, the five networks trained in the LOOCV folds were combined by averaging their output probability maps, and performance metrics were computed directly on these averaged probabilities. This averaging reduces variance and provides a more robust estimate of generalization performance on the external patient.

The primary objective is to minimize a weighted Dice loss function, which explicitly addresses the strong class imbalance inherent in this segmentation task. The Dice loss is defined in Equation (1), where $p$ is the predicted mask, $g$ is the ground-truth mask, and $\epsilon$ is a small constant to ensure numerical stability. By directly maximizing the overlap between predicted and ground-truth masks, Dice loss ensures that small but clinically critical structures such as perforators contribute proportionally to the optimization process.

$$Dice\ loss = 1 - \frac{2\sum_i p_i g_i + \epsilon}{\sum_i p_i + \sum_i g_i + \epsilon} \quad (1)$$

To further emphasize perforator detection, class weights of 0.8 (abdomen), 0.8 (umbilicus), 6.0 (perforators), and 2.4 (creases/artifacts) were applied. These weights were chosen to increase the relative importance of perforators and, to a lesser extent, creases, ensuring that the networks focus on structures most relevant for clinical decision-making. The weighted contributions are summed to compute the final weighted dice loss. Both training loss and validation loss were recorded during each epoch, and their convergence behavior was analyzed to compare the learning efficiency and generalization capabilities of the two models.

Both networks were initialized using PyTorch's default random weight initialization scheme. The models were trained from scratch, without the use of pretrained encoders, backbones, or externally sourced weights. Hyperparameters were fixed a priori, guided by established practices in deep learning and medical image segmentation. We used the Adam optimizer with learning rate 0.001, consistent with the values demonstrated to work well in practice [27]. A small batch size of 16 was chosen to improve generalization performance and limit memory requirements, aligning with findings that small mini-batches (typically 16–32) often yield better results in CNN training [28]. Maximum training was set to 100 epochs, with early stopping if validation loss did not improve for 10 consecutive epochs, a protocol frequently used in benchmark medical imaging studies [29], [30].

Two neural network architectures are evaluated and compared: the U-Net proposed by Ronneberger et al. [24] and a modified U-Net architecture developed by De La Hoz et al. [22].

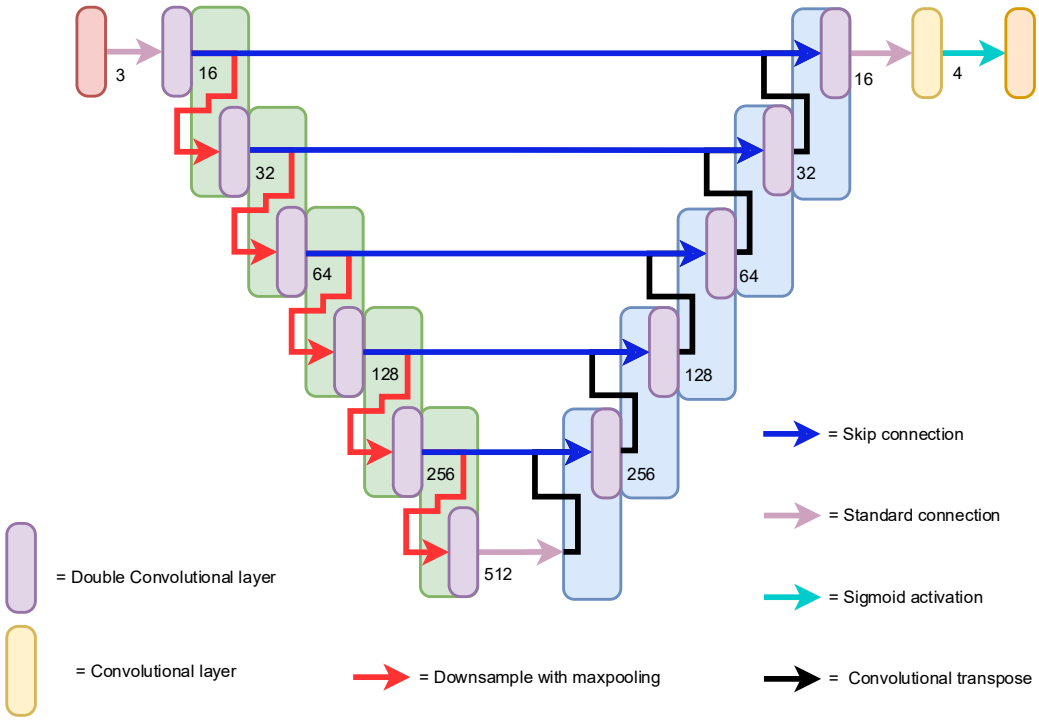



*Figure 3. The U-Net architecture with a symmetric encoder-decoder structure and skip connections.*

- The U-Net, introduced by Ronneberger et al. [31], is a widely used convolutional neural network designed for biomedical image segmentation. Its architecture follows a symmetric encoder-decoder structure, as illustrated in Figure 3. The encoder progressively downsamples the input image by applying 3x3 convolutions followed by ReLU activations and 2x2 max pooling operations. This process allows the network to capture multi-scale contextual information, which is critical for precise segmentation. The decoder, on the other hand, upsamples the feature maps back to the original resolution using transposed convolutions and combines them with high-resolution features from the encoder via skip connections. These skip connections help preserve fine-grained spatial information, which is crucial for accurately segmenting small, detailed structures such as perforators.

  The final layer of the U-Net consists of a 1x1 convolutional layer that produces a probability map with four channels, corresponding to the four segmentation classes: cooled abdomen, umbilicus, perforators, and cooling artifacts. U-Net's strength lies in its ability to generalize well, even when trained on relatively small datasets, making it particularly suitable for applications involving limited data availability, such as this study.

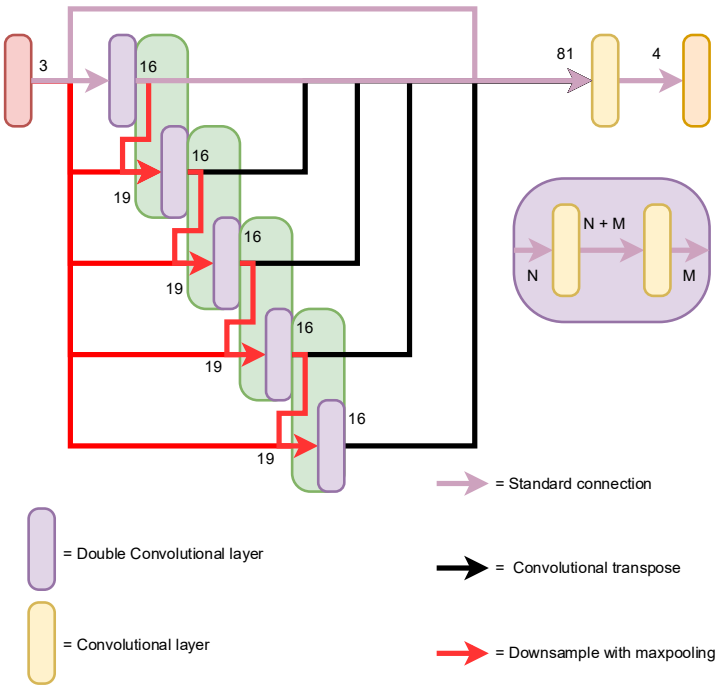



*Figure 4. The modified U-Net architecture by De La Hoz et al. [20], featuring a feature pyramid network design and 1x1 transformer layers for efficient upsampling.*

- The second network is a modified version of U-Net proposed by De La Hoz et al. [22], as illustrated in Figure 4. This model introduces key architectural changes aimed at improving computational efficiency and segmentation accuracy. Unlike the original U-Net, this architecture incorporates a feature pyramid network (FPN) design, where downsampling is applied at each convolutional block to process the input image at multiple resolutions. Each block receives a downsampled version of the original image, concatenated with the output from the previous block. This multi-resolution approach allows the network to capture features at various scales, which is particularly useful for detecting perforators of varying sizes.

  Additionally, instead of using transposed convolutions for upsampling, the modified U-Net employs a 1x1 transformer layer that directly processes resized block outputs. This change reduces computational overhead while maintaining high segmentation accuracy. The outputs of all blocks are upsampled to the original input size and concatenated before being passed to a final decision layer, which consists of a 1x1 convolution that generates the output segmentation map with four channels, corresponding to the four defined classes.

## 3. Results

In this section, we present the results of the neural network comparison in terms of training and validation weighted Dice loss, as well as segmentation performance. The evaluation is performed using patient-level LOOCV, with each iteration corresponding to a different patient being used as the validation set. This approach allows for the assessment of patient-specific variability and the robustness of the models under different conditions. In addition, an independent sixth patient was evaluated as an external test case to probe generalization beyond the cross-validation folds.

Figure 5 displays the training and validation weighted dice loss curves over 35 epochs for the U-Net and mU-Net models. The choice of 35 epochs aligns with findings that training for a fixed number of epochs, combined with early stopping, prevents overfitting and degradation of performance on validation data. Both networks show a steady decrease in training loss, with the U-Net achieving a lower final validation loss compared to the mU-Net. While the U-Net converges effectively, the mU-Net does not exhibit convergence, as its validation loss plateaus at a consistently higher value. These results indicate that the mU-Net struggles to generalize with the current measurements and settings, reflected in its inability to reduce validation loss effectively.

Table 1 reports the per-fold minima for training and validation weighted Dice losses for both models. For validation, the U-Net attains a mean loss of 0.42 ± 0.09, with a maximum inter-patient difference of 0.13, between validation

patients 3 and 4. While some folds exhibit slightly faster or slower convergence, the overall performance trends remain consistent across datasets. In contrast, the mU-Net exhibits substantially higher validation losses, with a mean of 0.87 ± 0.02, and no meaningful convergence across folds.

Segmentation performance was further quantified in terms of sensitivity and precision for perforator detection. Across the five LOOCV folds, U-Net achieved a mean sensitivity of 0.87 ± 0.08, indicating that most true perforators were consistently detected. The corresponding mean precision was 0.82 ± 0.14, reflecting the presence of occasional false positives. In the independent test patient, sensitivity remained high at 0.88, but precision dropped to 0.58, consistent with the larger number of false positives observed qualitatively.

These trends are further illustrated in Figure 6, which presents representative ground-truth labels and corresponding U-Net predictions. Only U-Net outputs are shown, as mU-Net produced near-uniform segmentations without meaningful anatomical detail. In all cases, the majority of perforators are correctly identified, confirming the robustness of the network. Occasional false positives and false negatives remain, highlighting opportunities for further optimization of the model. The independent test patient yielded a weighted Dice loss of 0.49, slightly higher than the LOOCV mean. Consistent with this, segmentation quality was generally preserved, although a greater number of false positives were observed compared to the cross-validation folds.

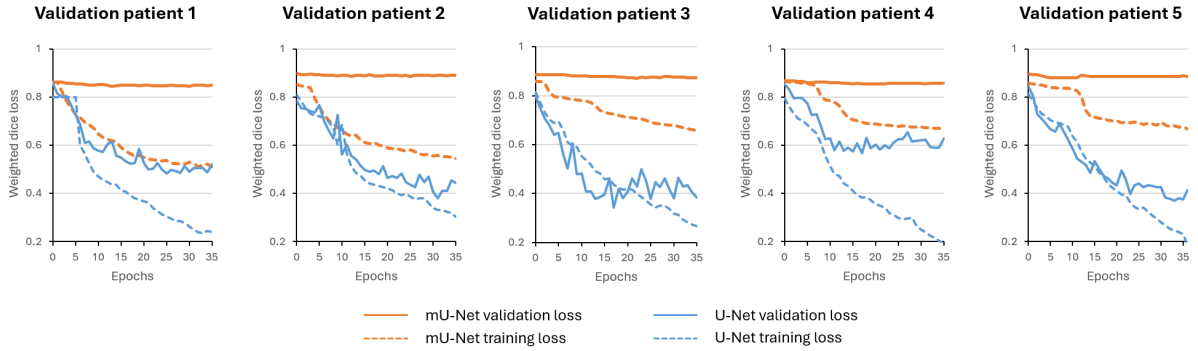



*Figure 5. The U-Net converges in all five folds, achieving a consistently lower validation loss, while the mU-Net fails to converge for any dataset.*

| | | Validation patient 1 | Validation patient 2 | Validation patient 3 | Validation patient 4 | Validation patient 5 |
|---|---|---|---|---|---|---|
| mU-Net | Training loss | 0.51 | 0.55 | 0.66 | 0.69 | 0.68 |
| | Validation loss | 0.84 | 0.89 | 0.87 | 0.85 | 0.88 |
| U-Net | Training loss | 0.24 | 0.30 | 0.27 | 0.19 | 0.23 |
| | Validation loss | 0.48 | 0.38 | 0.34 | 0.57 | 0.37 |

*Table 1. Training and validation weighted dice losses for both models after 35 epochs, using LOOCV where each patient is used as the validation set once.*

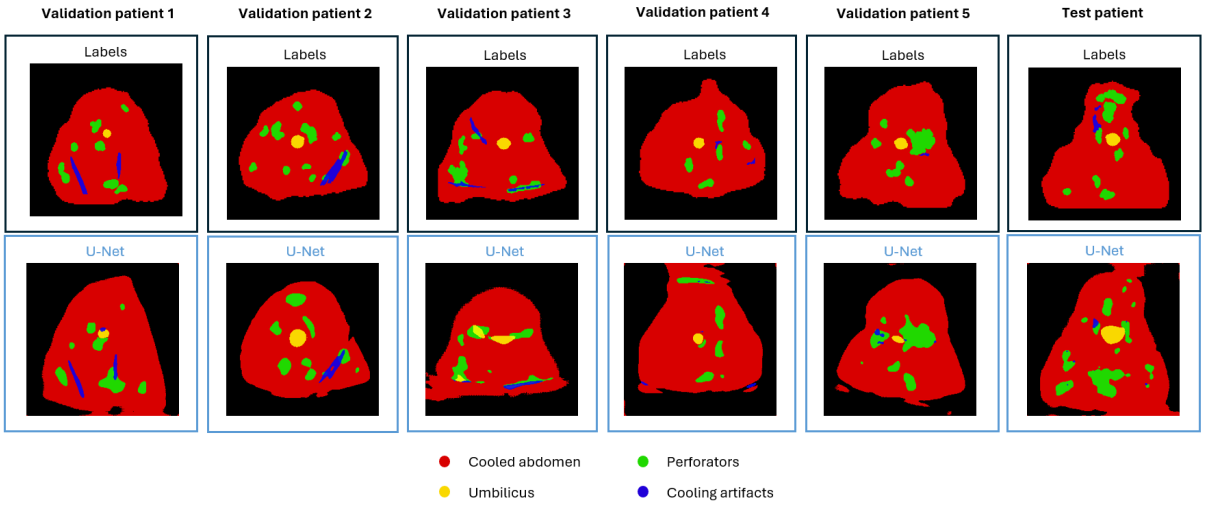


*Figure 6. Ground-truth labels (top) and U-Net predictions (bottom) for the five LOOCV validation patients and the test patient. The U-Net correctly identifies most perforators across patients, with boundary discrepancies and occasional false positives and false negatives, more pronounced in the test patient.*

## 4. Discussion

The results of this study demonstrate the effectiveness of the U-Net architecture for robust perforator detection in DIEP flap surgeries. Despite its relative simplicity, U-Net achieved superior generalization compared to the modified U-Net (mU-Net) proposed by De La Hoz et al. [22]. U-Net consistently reached lower validation losses across folds (mean weighted Dice loss 0.42 ± 0.09), with limited inter-patient variability, and maintained high sensitivity (0.87 ± 0.08) in detecting true perforators. Precision was somewhat lower (0.82 ± 0.14), reflecting the presence of occasional false positives, particularly in the independent test patient. These findings indicate that U-Net is capable of detecting clinically relevant perforators, although refinement is required to reduce false detections.

In contrast, the mU-Net failed to converge and produced near-uniform segmentations with persistently high validation losses (mean 0.87 ± 0.02). This pattern reflects underfitting and suggests that the architecture, under the present training conditions, is not capable of capturing the discriminative features required for perforator detection.

The robust performance of the U-Net can be attributed to its encoder-decoder structure with skip connections, enabling simultaneous capture of local and global contextual information. This facilitates detailed segmentation of perforators while maintaining robustness to inter-patient anatomical variability. However, segmentation errors at perforator boundaries and the occurrence of false positives underscore the need for further optimization to meet clinical standards.

This study is limited by the relatively small dataset of five patients, which, while diverse, may not fully capture the variability encountered in broader clinical populations. Additionally, the reliance on specific cooling protocols and imaging conditions may constrain the generalizability of our findings to other surgical setups or equipment configurations.

Future research should focus on expanding the dataset to include a more extensive range of patient anatomies, cooling methods, and surgical conditions. This would enhance the model's generalizability and robustness. In addition, hyperparameters were fixed a priori based on established practices rather than optimized for this specific dataset. While this ensured reproducibility and reduced computational cost, future work could systematically explore hyperparameter tuning to potentially enhance performance. Furthermore, optimization of the U-Net architecture, such as incorporating attention mechanisms or advanced loss functions, could further improve its segmentation accuracy. Lastly, integrating intraoperative validation could provide valuable insights into the real-time applicability of the model.

## 5. Conclusion

This study demonstrates that the U-Net architecture is a robust and effective tool for perforator detection in DIEP flap surgeries. By leveraging a diverse dataset and advanced preprocessing techniques, the U-Net achieved superior performance compared to the modified U-Net, highlighting its suitability for clinical applications. However, the observed segmentation errors and the dataset's limited size emphasize the need for continued research and refinement. Future work should aim to scale up data collection and optimize neural network architectures to improve accuracy.

**Author Contributions** Conceptualization, Warre Clarys; methodology, Warre Clarys; software, Rhys Evans and Warre Clarys; validation, Rhys Evans and Warre Clarys; formal analysis, Rhys Evans and Warre Clarys; investigation, Warre Clarys; resources, Warre Clarys; data curation, Warre Clarys; writing: original draft preparation, Warre Clarys; writing: review and editing, Warre Clarys, Rhys Evans, Jan Verstockt, Simon Verspeek, Hai Zhang; visualization, Rhys Evans and Warre Clarys; supervision, Veronique Verhoeven, Wiebren Tjalma, Filip Thiessen and Gunther Steenackers. All authors have read and agreed to the published version of the manuscript.

**Funding** This research was funded by the Research Foundation – Flanders (FWO) [https://doi.org/10.13039/501100003130] under a Strategic Basic Research (SB) fellowship [1SB9825N] (Warre Clarys), and by the Research Foundation – Flanders (FWO) [https://doi.org/10.13039/501100003130] through a grant for Applied Biomedical Research with a Primary Social Goal [FWO T001723N].

**Consent** The principles outlined in the Declaration of Helsinki have been followed and written informed consent was obtained from all patients for publication of this study and the accompanying images.

**Ethical Approval** the authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. Ethical approval has been obtained from ethics committee. Belgian registration number: B3002024000116.

**Disclosures** No potential conflict of interest was reported by the author(s).

**Abbreviations**

| | |
|---|---|
| CTA | Computed Tomography Angiography |
| CDU | Color Doppler Ultrasonography |
| DIRT | Dynamic Infrared Thermography |
| DIEP | Deep Inferior Epigastric artery Perforator |
| IR | Infrared |
| LOOCV | Leave-One-Out Cross-Validation |
| MRA | Magnetic Resonance Angiography |
| mU-Net | Modified U-Net |
| NETD | Noise Equivalent Temperature Difference |
| ReLU | Rectified Linear Unit |
| U-Net | Convolutional Neural Network for biomedical image segmentation |

**References**


[1] "Breast Cancer Facts and Statistics 2024." Accessed: Apr. 05, 2024. [Online]. Available: https://www.breastcancer.org/facts-statistics

[2] World Health Organisation (WHO), "Estimated number of new cases from 2022 to 2045, Both sexes, age [0-85+]," *Global Health Estimates 2022: Deaths by Cause, Age, Sex, by Country and by Region*, 2022, Accessed: Apr. 04, 2024. [Online]. Available: https://gco.iarc.fr/en

[3] T. T. Quinn, G. S. Miller, M. Rostek, M. S. Cabalag, W. M. Rozen, and D. J. Hunter-Smith, "Prosthetic breast reconstruction: Indications and update," Apr. 01, 2016, *AME Publishing Company*. doi: 10.3978/j.issn.2227-684X.2015.07.01.

[4] S. Epstein, B. N. Tran, J. B. Cohen, S. J. Lin, D. Singhal, and B. T. Lee, "Racial disparities in postmastectomy breast reconstruction: National trends in utilization from 2005 to 2014," *Cancer*, vol. 124, no. 13, pp. 2774–2784, Jul. 2018, doi: 10.1002/cncr.31395.

[5] American Society of Plastic Surgeons, "2022 ASPS Procedural Statistics Release," 2022. Accessed: Apr. 04, 2024. [Online]. Available: https://www.plasticsurgery.org/documents/News/Statistics/2022/plastic-surgery-statistics-report-2022.pdf

[6] L. Simion *et al.*, "Breast Reconstruction following Mastectomy for Breast Cancer or Prophylactic Mastectomy: Therapeutic Options and Results," *Life*, vol. 14, no. 1, p. 138, Jan. 2024, doi: 10.3390/life14010138.

[7] P. Sarkar *et al.*, "Rates of breast reconstruction uptake and attitudes toward breast cancer and survivorship among south asians: A literature review," Apr. 01, 2024, *John Wiley and Sons Inc*. doi: 10.1002/jso.27584.

[8] P. L. Myers, J. A. Nelson, and R. J. Allen, "Alternative flaps in autologous breast reconstruction," Jan. 01, 2021, *AME Publishing Company*. doi: 10.21037/gs.2020.03.16.

[9] A. Marwah *et al.*, "Outcomes of Deep Inferior Epigastric Artery Perforator (DIEP) Flap in Indian Population-A Prospective Single-Institute Study," *Indian Journal of Plastic Surgery*, 2024, doi: 10.1055/s-0044-1781445.

[10] F. E. F. Thiessen *et al.*, "Dynamic InfraRed Thermography (DIRT) in DIEP-flap breast reconstruction: A review of the literature," Nov. 01, 2019, *Elsevier Ireland Ltd*. doi: 10.1016/j.ejogrb.2019.08.008.

[11] R. G. Wade, J. Watford, J. C. R. Wormald, R. J. Bramhall, and A. Figus, "Perforator mapping reduces the operative time of DIEP flap breast reconstruction: A systematic review and meta-analysis of preoperative ultrasound, computed tomography and magnetic resonance angiography," Apr. 01, 2018, *Churchill Livingstone*. doi: 10.1016/j.bjps.2017.12.012.

[12] S. I. Moos, D. N. H. Van Vemde, J. Stoker, and S. Bipat, "Contrast induced nephropathy in patients undergoing intravenous (IV) contrast enhanced computed tomography (CECT) and the relationship with risk factors: A meta-analysis," Sep. 2013. doi: 10.1016/j.ejrad.2013.04.029.

[13] F. Zhao *et al.*, "Comparative effect of iso-osmolar versus low-osmolar contrast media on the incidence of contrast-induced acute kidney injury in diabetic patients: A systematic review and meta-analysis," *Cancer Imaging*, vol. 19, no. 1, Jun. 2019, doi: 10.1186/s40644-019-0224-6.

[14] W. J. Theuvenet, G. F. Koeyers, and M. H. M. Borghouts, "Thermographic assessment of perforating arteries," *Scand J Plast Reconstr Surg Hand Surg*, vol. 20, no. 1, pp. 25–29, 1986, doi: 10.3109/02844318609006287.

[15] S. E. Godoy *et al.*, “Dynamic infrared imaging for skin cancer screening,” *Infrared Phys Technol*, vol. 70, pp. 147–152, May 2015, doi: 10.1016/j.infrared.2014.09.017.

[16] M. Moderhak, Sz. Kołacz, J. Jankau, and T. Juchniewicz, “An Active Dynamic Thermography (ADT) method for TRAM flap blood perfusion mapping in breast reconstruction,” QIRT Council, Jul. 2016. doi: 10.21611/qirt.2016.089.

[17] W. Clarys, S. Verspeek, J. Verstockt, S. Hummelink, G. Steenackers, and F. Thiessen, “Comparative study of cooling techniques for perforator detection in DIEP flap reconstruction using dynamic infrared thermography,” *Journal of Plastic, Reconstructive & Aesthetic Surgery*, Aug. 2025, doi: 10.1016/j.bjps.2025.08.018.

[18] G. Steenackers *et al.*, “Infrared Thermography for DIEP Flap Breast Reconstruction Part II: Analysis of the Results,” MDPI AG, Dec. 2019, p. 49. doi: 10.3390/proceedings2019027049.

[19] L. De Weerd, S. Weum, and J. B. Mercer, “The value of dynamic infrared thermography (DIRT) in perforator selection and planning of free DIEP flaps,” *Ann Plast Surg*, vol. 63, no. 3, pp. 274–279, Sep. 2009, doi: 10.1097/SAP.0b013e318190321e.

[20] C. Mavioso *et al.*, “Automatic detection of perforators for microsurgical reconstruction,” *Breast*, vol. 50, pp. 19–24, Apr. 2020, doi: 10.1016/j.breast.2020.01.001.

[21] S. T. Kakileti and K. Venkataramani, “Automated Blood Vessel Extraction in TwoDimensional Breast Thermography,” IEEE, Aug. 2016. doi: 10.1109/ICIP.2016.7532383.

[22] E. C. De La Hoz *et al.*, “Automated thermographic detection of blood vessels for DIEP flap reconstructive surgery,” *Int J Comput Assist Radiol Surg*, Jul. 2024, doi: 10.1007/s11548-024-03199-8.

[23] Y. A. Ayalew, K. A. Fante, and M. A. Mohammed, “Modified U-Net for liver cancer segmentation from computed tomography images with a new class balancing method,” *BMC Biomed Eng*, vol. 3, no. 1, Dec. 2021, doi: 10.1186/s42490-021-00050-y.

[24] C. Li, Z. Li, F. Yu, and W. Liu, “An improved method for retinal vessel segmentation in U-Net,” *Multimed Tools Appl*, Oct. 2024, doi: 10.1007/s11042-024-18757-0.

[25] F. E. F. Thiessen *et al.*, “Dynamic infrared thermography (DIRT) in Deep Inferior Epigastric Perforator (DIEP) flap breast reconstruction: Standardization of the measurement set-up,” 2019, *AME Publishing Company*. doi: 10.21037/gs.2019.12.09.

[26] J. Verstockt, F. Thiessen, B. Cloostermans, W. Tjalma, and G. Steenackers, “DIEP flap breast reconstructions: thermographic assistance as a possibility for perforator mapping and improvement of DIEP flap quality,” *Appl Opt*, vol. 59, no. 17, p. E48, Jun. 2020, doi: 10.1364/ao.388351.

[27] D. P. Kingma and J. L. Ba, “Adam: A Method for Stochastic Optimization,” *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, Dec. 2014, Accessed: Aug. 12, 2025. [Online]. Available: https://arxiv.org/pdf/1412.6980

[28] N. S. Keskar, J. Nocedal, P. T. P. Tang, D. Mudigere, and M. Smelyanskiy, “On Large-Batch Training for Deep Learning: Generalization Gap and Sharp Minima,” *5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings*, Sep. 2016, Accessed: Aug. 12, 2025. [Online]. Available: https://arxiv.org/pdf/1609.04836

[29] J. Yang *et al.*, “MedMNIST v2 - A large-scale lightweight benchmark for 2D and 3D biomedical image classification,” *Sci Data*, vol. 10, no. 1, pp. 1–10, Dec. 2023, doi: 10.1038/S41597-022-01721-8;SUBJMETA=114,129,1305,631,648,697,706;KWRD=DATABASES,MACHINE+LEARNING.

[30] V. B. Sadu *et al.*, “Optimizing the early diagnosis of neurological disorders through the application of machine learning for predictive analytics in medical imaging,” *Sci Rep*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-05888-z.

[31] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” May 2015, [Online]. Available: http://arxiv.org/abs/1505.04597