IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2024.Doi Number

Can deep learning wound segmentation algorithms developed for a dataset be effective for another dataset? A specific focus on diabetic foot ulcers

Stuardo Lucho[1], Roozbeh Naemi [2,3], Benjamin Castañeda [1], Sylvie Treuillet [4]

[1] Departamento de Ingeniería, Pontificia Universidad Católica del Perú, 15088 Lima, Perú
[2] Centre for Human Movement and Rehabilitation, School of Health and Society, University of Salford, Manchester, M6 6PU, UK
[3] School of Health, Science and Wellbeing, Staffordshire University, Stoke-on-Trent, United Kingdom ST4 2DF
[4] Laboratoire PRISME, Université d'Orléans, 45067 Orléans, France

Corresponding author: Stuardo Lucho (e-mail: stuardo.lucho@pucp.edu.pe).

ABSTRACT Diabetic foot ulcers (DFU) represent a severe complication, often resulting from poor glycemic control, neuropathy, peripheral vascular disease, or inadequate foot care. DFUs can lead to significant morbidity, including amputation and, in severe cases, can be fatal. Recently, advancements in computer vision technologies based on artificial intelligence (AI) have shown promise in DFU management. Particularly deep learning (DL) models such as U-Net and other models and techniques, were utilized to enhance wound segmentation accuracy. This research focuses on evaluating the generalization capabilities of DL models across different DFU datasets. Specifically, we investigated whether models trained on one dataset can be effective when utilized on another dataset, addressing the challenge of cross-dataset generalization. We employed 7 popular DL models, U-Net-VGG16, U-Net-EfficientNetV2S, ABANet, MA-Net, LinkNet, DeepLabV3+, and Segment Anything Model (SAM), with 2 DFU datasets: FUSeg challenge and DFUC challenge. A total of 54 experiments were conducted plus 27 for SAM, involving training on one dataset, and testing on another, as well as training and testing on combined datasets. The results indicate substantial variability in segmentation performance when models trained on one dataset are tested on another, highlighting the influence of dataset characteristics on model generalization. The study underscores the importance of using diverse and comprehensive datasets to develop robust DL models for DFU segmentation and its generalization. This research contributes to the understanding of DL model performance in medical image segmentation and emphasizes the need for standardized datasets in improving DFU management through computer vision.

INDEX TERMS Wound segmentation, deep learning, Diabetic foot ulcers

I. INTRODUCTION

Diabetic foot ulcers (DFU) are ones of the most common and serious complications of diabetes mellitus, commonly caused by poor glycemic control, underlying neuropathy, peripheral vascular disease, or poor foot care [1]. They are characterized by open sores or wounds, typically located on the feet, and without the proper care (applying an interprofessional approach), can lead to significant morbidity, in severe cases, amputation, and in the worst scenario, and can be fatal for 50% of patients in a period of 5 years [2].

Recently, computer vision technologies based on artificial intelligence (AI), have been used for wound analysis [3], [4], [5], and also in the context of DFU. Deep learning (DL), which are neural networks with multiple layers (also known as deep neural networks or DNNs), such as U-Net have been used as a basis for many studies in wound segmentation [6], [7], which is essential for effective diagnosis and treatment planning. In Niri [8], the authors

change the encoder and use transfer learning to use U-Net-VGG16; in Mahbod [9], they used U-Net-EfficientNet and LinkNet for wound segmentation; and other models such as Mask R-CNN [10], post-processing techniques [11], or in some cases custom architectures [12], [13]. Likewise, the work by Bouallal [14] for early wound detection or the research by Gupta [15], for quantifying the extent and depth of the ulcer, which aids to monitor the progression of the wound over the time through wound segmentation. To aid in the detection of the different types of wound tissues (necrotic, sloughy, granulating and epithelializing) by Wang[16], the work by Niri [17], in the 3D analysis of DFU based on color and thermal photos [8], or for wound characteristics, such as wound bed, peri wound and normal skin, by Gutierrez [18].

One initiative to develop and improve segmentation methods for DFU is the "Diabetic Foot Ulcer Challenges" (DFUC). These challenges were launched in 2020 by the Medical Image Computing and Computer Assisted Interventions Society (MICCAI), as an annual international conference where different teams from around the world try to achieve the highest quality segmentation masks. The metric use to compare the results and declare the winner team of the challenge is the F1-score, also known as Dice (see Appendix A).

Training a robust segmentation model requires a well-curated images dataset that includes diverse types of DFU. Annotated images with ground truth, also called masked images or labeled images, are crucial for training and evaluating the performance of the segmentation algorithm. Unlike other datasets such as ImageNet, which has 1 281 167 generic images, diabetic foot ulcer-only datasets have much fewer images, with the three largest ones available upon request shown in Table I. While there are other datasets used in many research studies, they are either too small (less than 250 images) [19], [20] or made from chronic wounds (not only diabetic foot ulcers) [21].

Despite this, all previous studies have used the different DFU datasets for training (and the same for testing), but there is no study that use cross-datasets for training and testing. This means, while a DL model can get really good results and metrics for one dataset, such results may not be repeated when it is used in another DFU dataset. The question raised was whether it is possible to generalize a DL model trained on one dataset to predict semantic segmentation on images of a different dataset.

TABLE I
SUMMARY OF THE THREE LARGEST DFU PUBLIC DATASETS

Dataset	Resolution	Total images
AZH Wound Care Center Dataset [22]	224 x 224	1109
FUSeg 2021 Dataset (FUSeg The Foot Ulcer Segmentation Challenge) [23]	512 x 512	1210
DFUC 2020 Dataset (Diabetic Foot Ulcer Challenge 2020) [24]	640 x 480	4000

The purpose of this research is to use 7 DL models used in the literature for segmentation, train them on one dataset and test them on the other dataset and vice versa. Then, training and testing will be performed by combining both datasets and analyzing the results. The goal was not to develop a new DL model or architecture for DFU, but to make a qualitative comparison of the influence of the datasets on training, testing, and generalization of DL models for DFU segmentation.

The rest of this research is organized as follows. Section 2 presents the related work on DFU segmentation; Section 3 shows the deep learning models used, information about the datasets and the experiments performed; Section 4, describes the results and discussion over the information gated, and finally, Section 5 the conclusions.

II. RELATED WORK

Over the years, numerous research have explored several deep learning architectures to improve DFU segmentation performance, with the goal of achieving the highest metrics along with the most accurate definition of the wound.

In the Diabetic Foot Ulcers Grand Challenge (DFUGC) [25], several research teams propose novel architectures to get better DFU segmentation. The dataset used was composed of 4000 images (for training and testing) with their corresponding ground truth (GT), having an image resolution of 640x480px. In the most recent report from the DFUGC 2022 [26], the authors detailed the 5 top architecture with the highest Dice achieved during the challenge: 1) HarDNet-DFUS [27] (72.87%); 2) OCRNet with Edge Loss [28] (72.80%); 3) An optimal combination of BCE and Dice Losses with OoD [29] (72.63%); 4) A join model of transformers and CNN [30] (72.54%); 5) An ensemble of Feature Pyramid Network with an SE-ResNeXt101-32x4d backbone [31] (72.20%). The different algorithms used vary between CCN, FPN, Vision Transformers (ViT) and even in one case, enriched with synthetic generated data to increase the train set [31].

Outside the DFUGC, in the study conducted by Toofanee [32], the authors propose DFU-SIAM, an ensemble of CCN and ViT within a Siamese Architecture, train on 5955 images of 4 categories to classify infection, ischemia, both and none of them. The dataset used was the Diabetic Foot Ulcer Challenge 2021 [33]. In toledo [34], the authors propose MsBNet, a neural network not only to segment the diabetic foot ulcer, but to classify different kind of tissues (healthy, granulation, slough and necrotic) inside the wound. For that research, a private dataset was used. In the same way, the work by Nagaraju [35], which introduces a novel architecture for DFU detection and classification, named SSODL-DFUDC, that was trained on a public dataset of 844 samples of Normal and abnormal wounds (diabetic foot ulcer wounds).

The work by Lan [13], proposes an architecture capable of not only detecting wounds, but distinguishing between chronic wounds and diabetic foot ulcers, offering valuable support for

less experienced doctors. Within the same area of research, the work conducted by Kuo [36] proposed a data augmentation technique for DFU, TransMix, which combines Augmented Global Pre-training and Localized CutMix Fine-tuning, to increase the training set, which is a key point to achieve better results.

Finally, in 2023, the Fundamental AI Research team (FAIR) from Facebook released the Segment Anything Model (SAM) [37], which is a new foundational model based on ViT, with zero-shot learning for general-purpose object segmentation. SAM also allows input prompts, as bounding boxes, points, or event text, for segmentation in specific areas of an image. The work proposed by Chen[38], where the HarDNet-MSEG model is used to segment the diabetic foot ulcer and SAM for segment all the image. Then they perform a pixel-wise ensemble between the DL predicted mask and the SAM mask, achieving better results. As so, the work by Taipe [39] follows the same general idea, by using YoloV5 to locate the DFU and then use bounding boxes of the previously detected wound as SAM prompt input to improve the segmentation accuracy.

III. METHODOLOGY

In this section, we describe the deep learning models used as reference, as well as the datasets used for the training and testing phases. We also detail the metrics used for evaluating the deep learning models and afterward, the experiments performed with cross train-test dataset used by both deep learning models.

A. DEEP LEARNING MODELS

A widely used DL model by different research projects in segmentation for DFU is U-Net [40], which is a convolutional neural network (CNN) architecture used for image segmentation tasks in Computer Vision and medical image processing, introduced in 2015.

The U-Net architecture, as shown in Fig. 1, is characterized by a U-shaped structure, which consists of a downward path “encoder”, an upward path “decoder” and a connection layer “bridge”. The encoder part of the U is used to capture the features of the image and reduce its spatial resolution. It consists of repeated two 3x3 convolutions, each followed by a rectified linear unit (ReLU) and a 2x2 max pooling operation. After the encoder, there is a "bridge" connecting the encoder to the decoder. This connection helps retain important spatial information that may be lost during the down sampling process.

The upward part of the U is called the decoder and is used to increase spatial resolution. It consists of an upsampling of the feature map followed by a 2x2 up-convolution channels, a concatenation with the correspondingly cropped feature map from the down sampling, and two 3x3 convolutions, each followed by a ReLU. At the last layer a 1x1 convolution is used to map each 64- component feature vector.

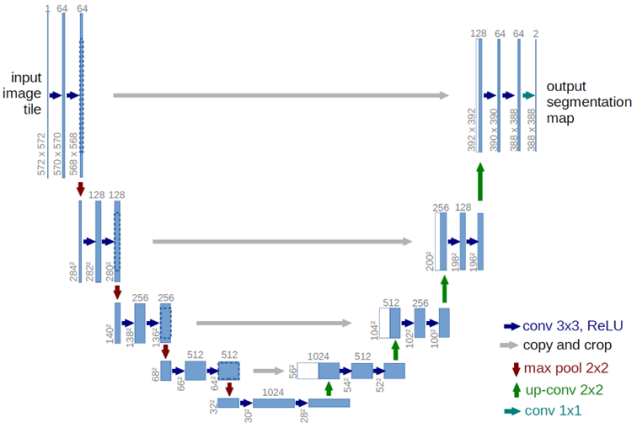


FIGURE 1. U-Net architecture (from [40]).

To avoid training a U-Net architecture from the scratch, transfer learning was used in this study. Transfer learning is a machine learning technique where a model trained on one task is adapted to work on a second one. As the Diabetic foot ulcers dataset are small compared to large datasets like ImageNet, transfer learning allows the model to benefit from the knowledge gained from the pretrained model. Also, by using transfer learning, it is possible to reduce the training time on the final model.

For these research, 7 DL architectures were tested, using transfer learning, as along with different encoders as follows:

1) U-NET - VGG16

VGG16 [41] is a convolutional neural network architecture characterized by its deep architecture with 16 weight layers (13 convolutional layers and 3 fully connected layers). VGG16 is often used for image classification tasks and is known as a feature extractor in various computer vision applications. Due to its ability to extract features, VGG16 has been placed as a U-NET encoder. The VGG16 model was pretrained using ImageNet.

2) U-NET - EFFICIENTNETV2S

EfficientNetV2S refers to a variant of the EfficientNetV2 model [42], which is an advanced convolutional neural network architecture designed for image classification and is known for its efficiency in terms of both accuracy and computational resources. The "S" in EfficientNetV2S denotes a smaller set of parameters (23.9 million) compared to “M” or “L”, small GMACs (4.9 giga) and small activation functions (21.4 million), generating a computationally less expensive algorithm while still maintaining good performance. Combining EfficientNetV2S with U-Net involves using the EfficientNetV2S as the encoder and integrating it within the U-Net architecture, where it takes the features extracted by EfficientNetV2S and refines them for the segmentation task. The skip connections between the encoder and decoder help maintain spatial information crucial for accurate wound

segmentation. The EfficientNetV2S model has also been pretrained using ImageNet.

3) ABANET

Proposed by Rezvani [43], ABANet is a novel attention-based network designed for precise image segmentation, focusing on masked face recognition during the COVID-19 pandemic. Even though the model is not made for DFU segmentation, it utilizes attention gates (AG) in a U-Net-based encoder-decoder architecture, which is a typical network used for DFU wounds. The model incorporates a hybrid loss combining focal, SSIM, and IoU losses to enhance boundary mask prediction.

4) MA-NET

A novel deep learning model designed for liver and tumor segmentation in CT images by Fan [44]. It enhances U-Net by incorporating a self-attention mechanism, focusing on both spatial and channel dependencies to improve segmentation accuracy. The model features two key components: the Position-wise Attention Block (PAB) for capturing spatial dependencies between pixels and the Multi-scale Fusion Attention Block (MFAB) for integrating multi-scale semantic information. It uses a combination of Dice and cross-entropy loss to improve performance. To improve the feature extraction, EfficientNet-b7 [45] was used as encoder pretrained on ImageNet.

5) LINKNET

LinkNet [46], is a DL architecture that uses an encoder-decoder structure inspired by auto-encoders, where encoder outputs are directly linked to corresponding decoders, bypassing layers to preserve spatial information and reduce computational load. LinkNet uses a lightweight ResNet18 backbone, offering significantly faster performance than alternatives like SegNet and ENet. For this model, EfficientNet-b7 was used as encoder pretrained on ImageNet.

6) DEEPLABV3+

DeepLabv3+ [47], an improve network over DeepLabv3, combines the strengths of spatial pyramid pooling and the encoder-decoder structure for semantic segmentation. DeepLabv3+ incorporates a decoder module to refine object boundaries, improving the segmentation accuracy. The authors utilize depth-wise separable convolution in both the Atrous Spatial Pyramid Pooling (ASPP) and decoder modules, resulting in faster and more efficient computation. For DeepLabv3+, after some initial tests, Resnet50 was used as encoder with ImageNet as transfer learning.

7) SAM POST-PROCESSING

Taking as reference the research papers presented in the "Related Work" section, an experiment with SAM was proposed as follows: First, the wound segmentation was performed with 3 DL networks (Ma-Net, LinkNet and DeepLabV3+) and then from the generated mask, a bounding box was obtained. This box was used as an input prompt for SAM, together with the original image, obtaining a segmentation in the same place but with the accuracy of SAM. This flow has its limitations, because if the original DL model fails to find a DFU wound, then SAM will not be able to find the wound either. A detailed flow of the use of SAM is shown in appendix B.

In the context of DL, the training loss function (also known as the cost function) represents a crucial component. This function serves as a metric for evaluating the accuracy of a model's predictions in comparison to the ground truth during the training phase. There are various types of loss functions, such Mean Squared Error, Huber Loss, Categorical Cross-Entropy, among others. The selection of the appropriate loss function depends on the specific task (classification, segmentation, etc.) and the nature of the data. The choice of the loss function has a significant impact on the performance of the DDN. In this research, two loss functions were selected.

- *Categorical Cross-Entropy (CCE) loss:*
  Employed in multiclass classification tasks, where the model predicts the probability distribution over multiple classes.
- *Dice Loss:*
  Commonly utilized in tasks such as semantic segmentation, it measures the overlap between the predicted mask and the ground truth, particularly useful to accurately delineate object boundaries.

B. IMAGE DATASET

To perform the experiments, 2 datasets were used.

1) FUSEG 2021 DATASET

FUSeg The Foot Ulcer Segmentation Challenge [23], Diabetic foot ulcer dataset that contains 1210 foot ulcer images from 889 patients. The ground truth for all the images were annotated by wound care experts and split into a training set (1010 images) and a testing set (200 images). The resolution of all the images is 512x512px (262,144 total). An example of a photo and its corresponding ground truth is shown in Fig. 2.

FIGURE 2. Example of FUSeg 2021 Dataset. (a) Original image (b) Ground truth.

2) DFUC 2020 DATASET

Diabetic Foot Ulcer Challenge 2020 [48], [49], [50]: which contains 4000 images used for training and testing. The ground truth was produced by two healthcare professionals (a podiatrist and a consultant physician) who specialize in diabetic wounds and ulcers. The resolution of all the images is 640x480px (307,200 total). An example of a photo and its corresponding ground truth is shown in Fig. 3.

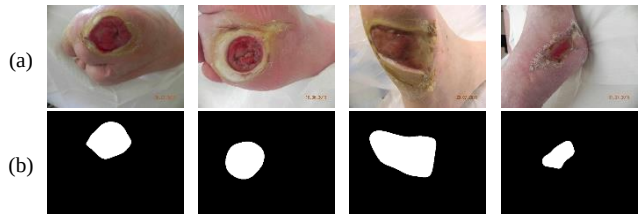


FIGURE 3. Example of DFUC 2020 Dataset. (a) Original image (b) Ground truth.

A summary of both datasets used is shown in Table II. For the case of DFUC dataset, only 2000 images had ground truth, so the train-test split will be based on this restriction; as so for FUSeg dataset, where only 1010 images had ground truth. In both cases, 80% of the images were used for training (and validation) and 20% for testing. We also considered a join of both datasets to test the results.

TABLE II
SUMMARY OF DATASETS USED

Dataset	Total images	Train images	Test images	Images resolution
FUSeg	1010	808	202	512 * 512 px 262,144 px
DFUC	2000	1600	400	640 * 480 px 307,200 px
DFUC + FUSeg	3010	2408	602	Mixed of FUSeg and DFUC

Using the ground truth for each photo, the ratio between the size of the wound and the size of the whole photo was calculated for both datasets, as shown in Fig. 4. The size of the wound compared to the whole photo is small for both datasets, 0.01 - 2% for FUSeg and 0.01 – 5% for DFUC.

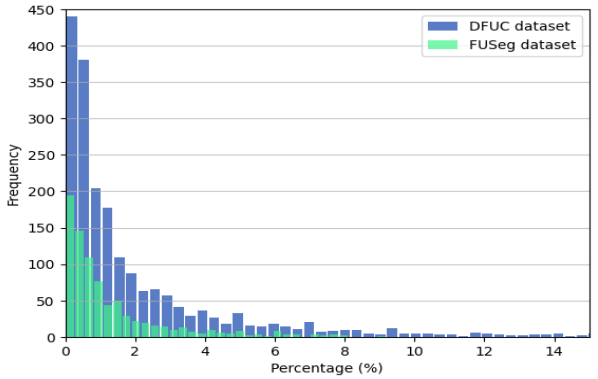


FIGURE 4. Wound-photo ratio histogram in percentage.

C. METRICS

IoU (Intersection over Union) and Dice coefficient (also known as F1 score) are evaluation metrics commonly used in image segmentation tasks and commonly reported indices for the algorithm efficacy of diabetic foot ulcer segmentation [51].

These metrics help assess the accuracy of the segmentation results by comparing the predicted segmentation masks with the ground truth masks. The details of the metrics are shown in appendix A.

D. PROPOSED EXPERIMENT

The experiments consisted of 6 DL models, as detailed in table III, along with the selected backbone.

TABLE III
DEEP LEARNING MODELS USED

Model	Backbone	Params (millions)
U-Net	VGG16	14M
U-Net	EfficientNetV2S	24M
ABANet	-	21M
Ma-Net	Efficientnet-B7	67M
LinkNet	Efficientnet-B7	67M
DeepLabV3+	Resnet50	23M

For each DL model, 9 combinations were proposed, as shown in table IV, giving a total of 54 trials performed (9 for each DL model).

TABLE IV
COMBINATIONS PERFORMED FOR EACH DL MODEL

		Testing Dataset		
		DFUC+FUSeg	DFUC	FUSeg
Training dataset	DFUC+ FUSeg	x	x	x
	DFUC	x	x	x
	FUSeg	x	x	x

Furthermore, we selected 3 models from Table III and used SAM after DL prediction (as shown in appendix B), to analyze whether the segmentation was improved or worsened. We annotated the models as follows:

- *Ma-Net – Efficientnet-b7 – SAM*
- *LinkNet – Efficientnet-b7 – SAM*
- *DeepLabV3+ - Resnet50 – SAM*

Following the combination defined in Table IV, a total of 27 trials were performed with SAM as post-processing.

All the experiments were developed in Kaggle with the following parameters: Python version 3.10, tensorflow version 2 with keras, 0.001 learning rate, 8 batch size, 200 epochs and GPU P100. To have homogeneous data, the FUSeg images were resized during training from 512x512 to 640x480 to match the DFUC images.

IV. RESULTS AND DISCUSSION

The results will be discussed in two parts:

- *The result for the 54 trials for the DL models,*
- *The results concern the 27 trails with SAM as postprocessing for segmentation.*

In Fig 5, the results for the 54 trials for the DL models are shown, where the best Dice score achieved in a heatmap table. All the results are based on Dice score, because, as shown in Appendix C, all the IoU scores calculated for all the trials were linear to the Dice Score.

From the results shown in Fig. 5, as general observations, ABANet consistently outperformed other models in all datasets, even when tested on FUSeg dataset, the difference between the best score and ABANet is between 1% to 2%. U-Net-VGG16 generally performed worse than other models, even with different train datasets and test datasets. Ma-Net and LinkNet showed similar performance across datasets.

The best result obtain in the trials was Ma-Net-EfficientNet-B7 trained on FUSeg and tested on FUSeg with 83,03%; while the worst result was U-Net-VGG16 trained on FUSeg and tested on DFUC with 39.90%. For both, the same train set (FUSeg) was used, but when test on FUSeg (202 images), the Dice 83,03% decreased almost half of it, to 39.90%, when tested on DFUC (400 images. One hypothesis could be the influence of the model, causing the Dice to decrease, so to analyze this effect, in Fig. 6, a comparison of all the Models trained on FUSeg is shown.

All the models trained on FUSeg performed very well on FUSeg dataset (over 79%), however, all of them decreased 30% points in Dice approximately when tested on DFUC and 25% when tested in DFUC + FUSeg. Based on this observation, the DL model posed a small influence in these results compared to the dataset used to train the models and test them.

Clearly, for FUSeg training and FUSeg testing, all DL models achieved high Dice scores, considering the test set was comprised of only 200 images, which proved the models were overfitting. In fact, the highest results of all tests (for all the datasets) were for FUSeg datasets, but when tested on other datasets, the Dice decreased in all the trials, in some cases, almost half the Dice.

			Test Dataset		
DL Model	Encoder	Train Dataset	DFUC	DFUC+FUSeg	FUSeg
ABANet	-	FUSeg	48.90 %	60.03 %	82.29 %
DeepLabV3+	ResNet50	FUSeg	40.52 %	53.58 %	79.69 %
LinkNet	EfficienteNet-B7	FUSeg	40.78 %	54.94 %	82.72 %
Ma-Net	EfficienteNet-B7	FUSeg	41.48 %	55.52 %	83.03 %
U-Net	EfficientNetV2S	FUSeg	48.53 %	59.00 %	79.71 %
U-Net	VGG16	FUSeg	39.90 %	49.65 %	81.43 %
ABANet	-	DFUC	72.17 %	75.31 %	81.58 %
DeepLabV3+	ResNet50	DFUC	68.08 %	70.61 %	75.67 %
LinkNet	EfficienteNet-B7	DFUC	71.41 %	75.06 %	82.09 %
Ma-Net	EfficienteNet-B7	DFUC	71.17 %	74.30 %	77.09 %
U-Net	EfficientNetV2S	DFUC	70.76 %	73.05 %	77.64 %
U-Net	VGG16	DFUC	57.72 %	62.99 %	73.52 %
ABANet	-	DFUC + FUSeg	72.12 %	72.23 %	72.46 %
DeepLabV3+	ResNet50	DFUC + FUSeg	67.47 %	65.33 %	61.42 %
LinkNet	EfficienteNet-B7	DFUC + FUSeg	70.86 %	67.83 %	61.67 %
Ma-Net	EfficienteNet-B7	DFUC + FUSeg	71.96 %	69.54 %	64.22 %
U-Net	EfficientNetV2S	DFUC + FUSeg	70.06 %	71.20 %	74.42 %
U-Net	VGG16	DFUC + FUSeg	67.57 %	62.13 %	56.18 %

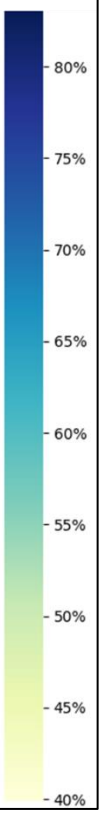


FIGURE 5. **Dice score for trials on deep learning models group by train set.**

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3502467

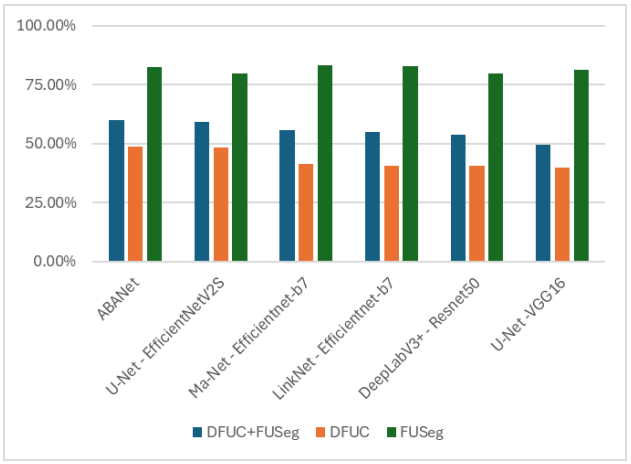


FIGURE 6. **Results of DL models trained on FUSeg**

In the same way, when the train set was DFUC and the mix DFUC + FUSeg, almost all the models performed very similarly, as shown in Fig. 7 and Fig. 8.

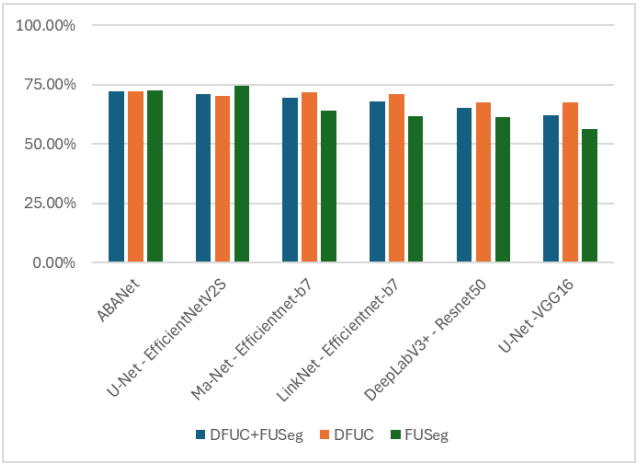


FIGURE 7. **Results of DL models trained on DFUC**

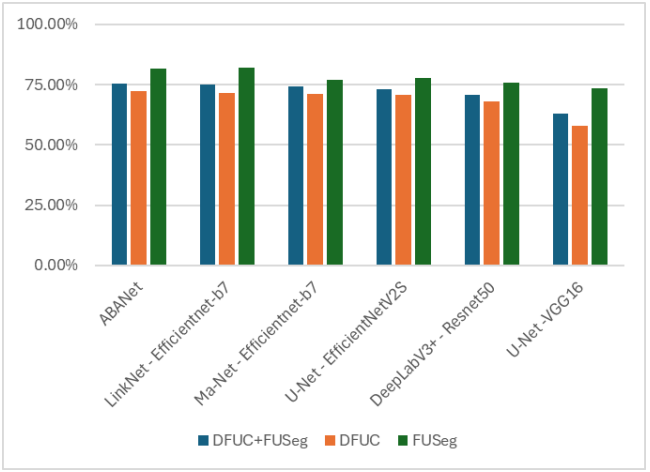


FIGURE 8. **Results of DL models trained on DFUC + FUSeg**

By comparing Fig. 7 and Fig. 8, when the train set is the merge of both datasets, the results increased in all cases for all DL models; however, these results were not as good as the ones trained on FUSeg and tested on FUSeg, meaning again, that the FUSeg dataset tends to overfits.

If we consider the model LinkNet trained on FUSeg, and then the one trained on DFUC, it stayed in 4th place; however, when trained on the combination of DFUC + FUSeg, it achieved 2nd place. Based on this observation, and taking in consideration that challenges rate the best algorithms based on the highest Dice achieved on certain dataset, if an algorithm perform average on some challenge, it could perform way better when trained on a wider dataset.

From Fig. 6, 7 and 8, ABANet and Unet-EfficientNetV2S were among the most generalizable across test sets, as they maintained high scores with minimal variance across the training datasets.

Another interesting result is the distribution of the Dice group by Model, as shown in Fig. 9.

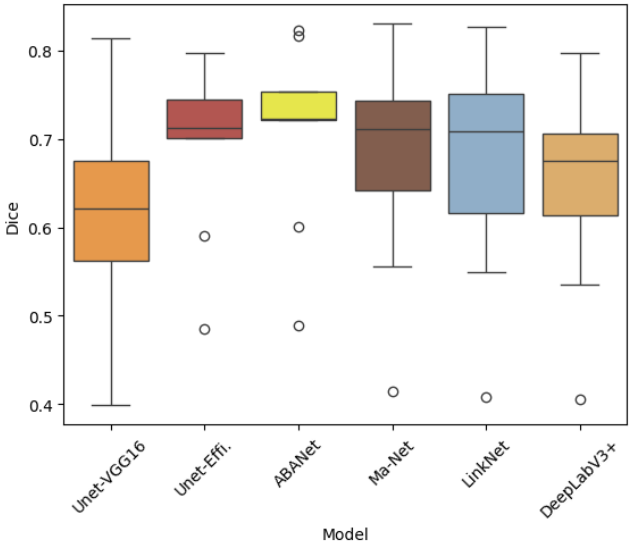


FIGURE 9. **Dice Score distribution group by model**

U-Net-VGG16 has the highest variability, indicating it performs less consistently across training datasets.

ABANet and U-Net-EfficientNetV2S maintain consistently higher scores across datasets, while DeepLabV3+, Ma-Net and LinkNet show lower and more variable performance.

Some examples of the comparison between the ground truth and the predicted images using different kinds of DL models, training and test sets for FUSeg and DFUC test sets is shown in Fig. 10.

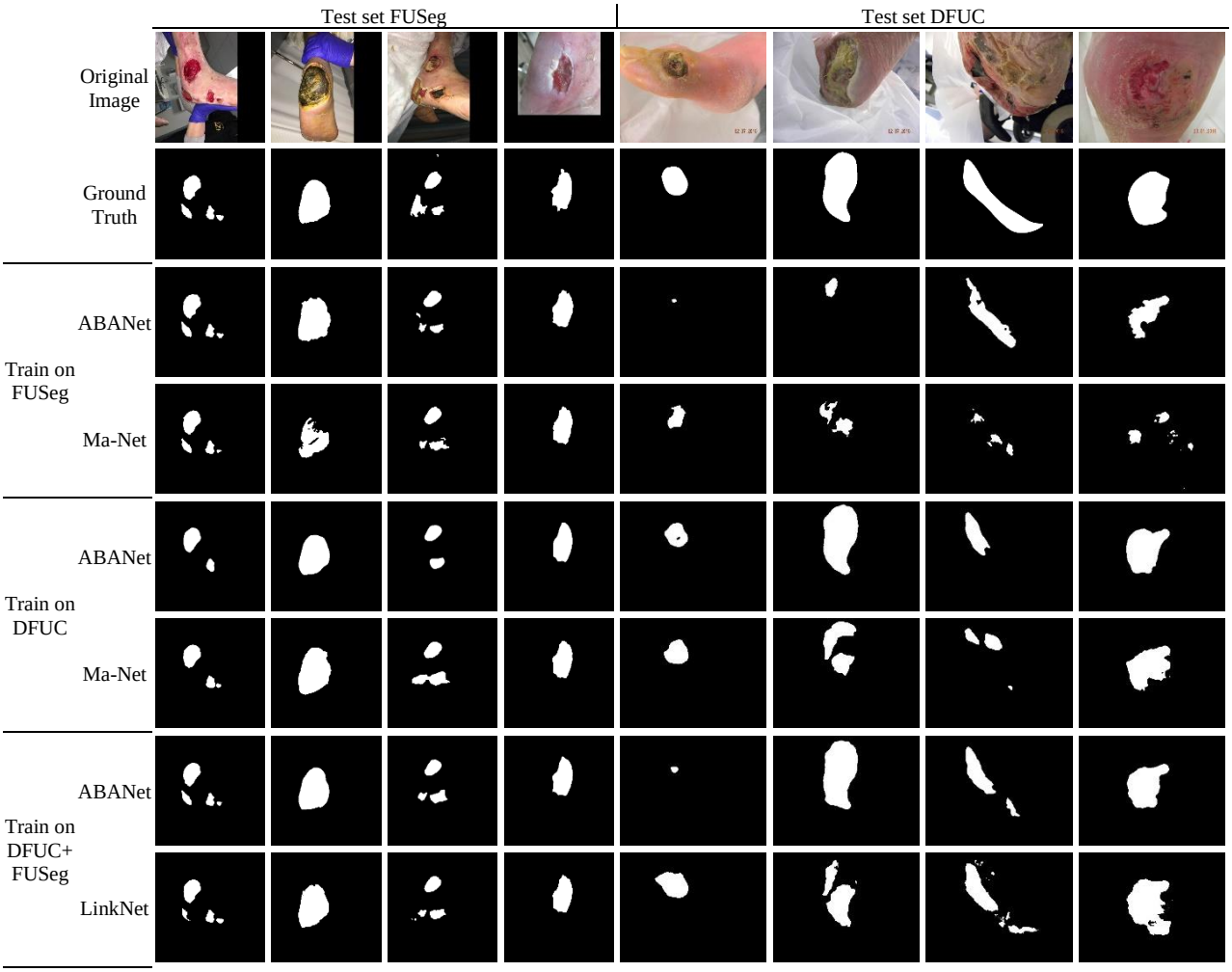


FIGURE 10. **For each train set, the 2 best results that were tested on each dataset.**

Regarding SAM, the 27 trials result for the 3 selected models trained on FUSeg, DFUC and DFUC + FUSeg using SAM as post-processing segmentation method (see appendix B), is shown in table V and Fig. 11.

TABLE V
RESULTS OF DL MODELS USING SAM AS POST-PROCESSING METHOD FOR SEGMENTATION

Model	Train dataset	Test set		
		DFUC+ FUSeg	DFUC	FUSeg
Ma-Net – Efficientnet-b7 - SAM	DFUC	70.92%	70.40%	71.44%
	FUSeg	58.15%	45.23%	83.73%
	DFUC+FUSeg	73.59%	69.19%	82.51%
LinkNet - Efficientnet-b7 - SAM	DFUC	68.12%	68.04%	67.58%
	FUSeg	55.34%	41.04%	83.41%
	DFUC+FUSeg	72.86%	67.70%	83.38%
DeepLabV3+ - Resnet50 - SAM	DFUC	67.00%	66.47%	67.92%
	FUSeg	56.50%	43.80%	81.90%
	DFUC+FUSeg	71.21%	67.35%	78.94%

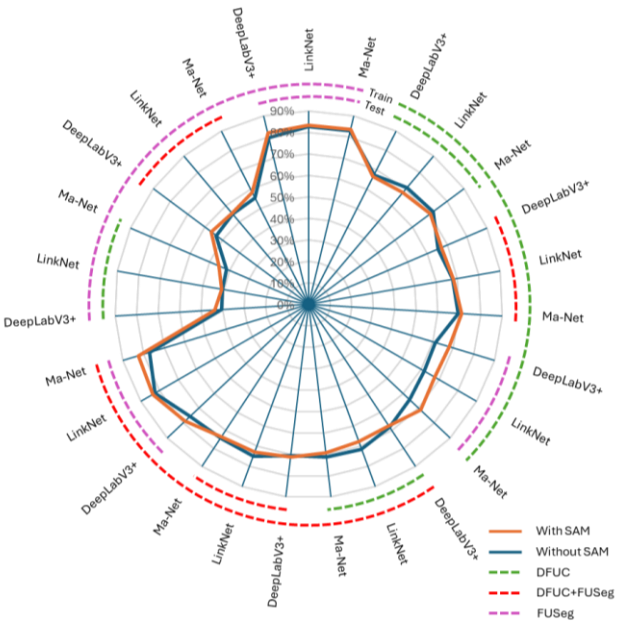


FIGURE 11. **Results of DL models trained on DFUC + FUSeg**

Based on table V, 70% (19 out of 27) of the results improved when using SAM as posprocessing method, with Dice score between 0.26% and 7.23%, while the other 8 decreased their Dice score between 0.71% and 3.71%.

A more comprenhensive analysis is shown in Table VI, where the biggest improvement for LinkNet was 5.91%, DeepLabV3+ with 6.51% and Ma-Net with 7.23%, when trained on DFUC and tested on FUSeg.

TABLE VI
EFFECT OF USING SAM AS POST-PROCESSING

Model	Quantity	Train dataset	Test dataset	Dice	Increase range (%)
All	9	FUSeg	All datasets	↑	0.26 to 3.75
All	3	DFUC	FUSeg	↑	5.91 to 7.23
All	3		DFUC	↓	-1.01 to -2.82
All	3		DFUC+FUSeg	↑	0.28 to 1.67
All	3	DFUC+FUSeg	FUSeg	↑	1.29 to 5.43
All	3		DFUC	↓	-0.73 to -3.71
LinkNet, Ma-Net	2		DFUC+FUSeg	↓	-0.71 to -2.19
DeepLab	1		DFUC+FUSeg	↑	0.60

When the training set was small (FUSeg), SAM improved the segmentation in all test sets. However, when the training set was bigger (DFUC or mixed) SAM improved the segmentation only on unseen images (when tested on FUSeg); but on images similar to the ones the DL model was trained on, SAM decreased the Dice; which means that on smaller datasets, SAM can help to generalize segmentation on unseen images.

V. CONCLUSION

This paper analyzes the influence of datasets on different deep learning models and their generalization. The study revealed a significant variation in the performance of deep learning models trained on one dataset (FUSeg) and tested on another one (DFUC). For instance, all the models trained on FUSeg achieved high Dice scores when tested on the FUSeg dataset, but they all showed a big decrease on Dice score when tested on the DFUC dataset. This indicates that models might be overfitting to the training dataset and failing to generalize across different datasets, taking in consideration that FUSeg dataset (200 images) compared to the DFUC dataset (400 images) likely contributed to the overfitting observed. A more diverse and larger dataset on diabetic foot ulcers exclusively could potentially enhance the DL model's ability to generalize and perform consistently across different datasets.

The DL model ABANet achieved the best segmentation in most of the train-test combination cases, while VGG16-UNet shows the lowest Dice when train on FUSeg and test on DFUC. However, in general all the models showed the same trend, to decrease when train on FUSeg and test on DFUC but improved when trained on DFUC (or the mixed) and test on DFUC. This consistency across different DL models suggests that the observed performance drop (or improve) is due to dataset differences rather than model architecture inefficiencies.

The use of SAM as prompt-based segmentation model increased the results when the training set was small (FUSeg - 200 images), which confirms the boost SAM can provide on datasets which few training samples.

While deep learning models show promise for the segmentation of diabetic foot ulcers, the study underscores the importance of dataset diversity, size, and the need for robust evaluation methods to ensure reliable and generalized performance across different clinical datasets.

APPENDIX A
METRICS DETAILS

- *Intersection over Union (IoU):*
  IoU is a measure of the overlap between the predicted segmentation mask and the ground truth mask. It is calculated as the intersection of the predicted mask and ground truth, divided by the union of these regions. The formula for IoU is defined in (1).

$$IoU = \frac{Area\ of\ intersection}{Area\ of\ union} \quad (1)$$

  IoU values range from 0 (no overlap) to 1 (perfect overlap) with higher IoU values indicate better segmentation accuracy. To get a final score, the IoU of each predicted mask and ground truth is calculated and the mean of all IoU is computed.

- *Dice Coefficient (F1 Score):*
  The Dice coefficient, also known as the F1 score, is another measure of the similarity between the predicted and ground truth segmentation masks. It is calculated as twice the intersection of the predicted mask and ground truth region divided by the sum of the areas of these regions. The formula for the Dice coefficient is defined in (2).

$$Dice = \frac{2 \times Area\ of\ intersection}{Area\ of\ predicted + Area\ of\ ground\ truth} \quad (2)$$

  Dice coefficient values also range from 0 to 1, with higher values indicating better segmentation performance.

APPENDIX B
SAM AS POST-PROCESSING AFTER DEEP LEARNING MODEL

We propose a simple architecture that will use the segmentation mask of a deep learning model, which has been trained on diabetic foot ulcer wounds, get the bounding box of the segmentation mask output and then use this as in input prompt for SAM. The flow is shown in Fig. 12.

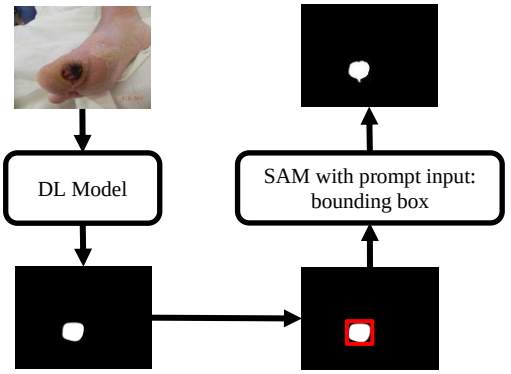


FIGURE 12. **SAM as post-processing for DL models.**

APPENDIX C
RELATION BETWEEN DICE AND IOU

For all the models tested, the relation between the Dice score and the IoU score was lineal.

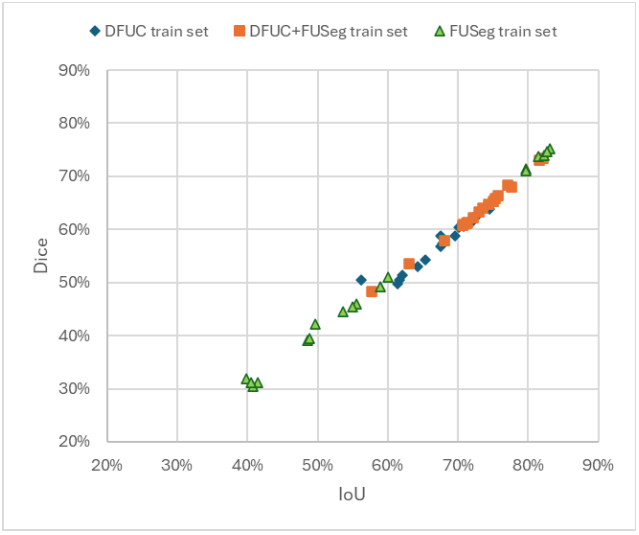


FIGURE 13. **Relation between Dice and IoU Score for all the trials**

ACKNOWLEDGMENT

The study was done as part of STANDUP Smartphone Thermal ANalysis for Diabetic foot Ulcer Prevention and treatment Project. The project is funded by the European Commission under Marie Skłodowska Curie Research and Innovation Staff Exchange (Horizon 2020 - MSCA - RISE - 2017, Grant Agreement Number 777661) Jan 2018 - Oct 2023.

REFERENCES

[1] W. J. Jeffcoate and K. G. Harding, "Diabetic foot ulcers," *The Lancet*, vol. 361, no. 9368, pp. 1545–1551, May 2003, doi: 10.1016/S0140-6736(03)13169-8.

[2] P. K. Moulik, R. Mtonga, and G. V. Gill, "Amputation and Mortality in New-Onset Diabetic Foot Ulcers Stratified by Etiology," *Diabetes Care*, vol. 26, no. 2, pp. 491–494, Feb. 2003, doi: 10.2337/diacare.26.2.491.

[3] R. Zhang, D. Tian, D. Xu, W. Qian, and Y. Yao, "A Survey of Wound Image Analysis Using Deep Learning: Classification, Detection, and Segmentation," *IEEE Access*, vol. 10, pp. 79502–79515, 2022, doi: 10.1109/ACCESS.2022.3194529.

[4] Y. Cao and Y. Wang, "The Impact of Artificial Intelligence and Deep Learning-Based Family-Centered Care Interventions on the Healing of Chronic Lower Limb Wounds in Children," *IEEE Access*, vol. 12, pp. 125557–125570, 2024, doi: 10.1109/ACCESS.2024.3454769.

[5] B. K. S. Kumar, K. C. Anandakrishan, M. Sumant, and S. Jayaraman, "Wound Care: Wound Management System," *IEEE Access*, vol. 11, pp. 45301–45312, 2023, doi: 10.1109/ACCESS.2023.3271011.

[6] V. Godeiro, J. S. Neto, B. Carvalho, B. Santana, J. Ferraz, and R. Gama, "Chronic wound tissue classification using convolutional networks and color space reduction," in *2018 IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP)*, 2018, pp. 1–6. doi: 10.1109/MLSP.2018.8517026.

[7] C. Pathompatai, R. Kanawong, and P. Taeprasartsit, "Region-Focus Training: Boosting Accuracy for Deep-Learning Image Segmentation," in *2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, Chonburi, Thailand: IEEE, Jul. 2019, pp. 319–323. doi: 10.1109/JCSSE.2019.8864162.

[8] R. Niri *et al.*, "Multi-View Data Augmentation to Improve Wound Segmentation on 3D Surface Model by Deep Learning," *IEEE Access*, vol. 9, pp. 157628–157638, 2021, doi: 10.1109/ACCESS.2021.3130784.

[9] A. Mahbod, G. Schaefer, R. Ecker, and I. Ellinger, "Automatic Foot Ulcer Segmentation Using an Ensemble of Convolutional Neural Networks," presented at the 2022 26th International Conference on Pattern Recognition (ICPR), IEEE Computer Society, Aug. 2022, pp. 4358–4364. doi: 10.1109/ICPR56361.2022.9956253.

[10] M. Pl, R. R, and M. N, "Automatic Segmentation of Diabetic foot ulcer from Mask Region-Based Convolutional Neural Networks," *J. biomed. res. clin. investig.*, vol. 2, no. 1, Aug. 2020, doi: 10.31546/2633-8653.1006.

[11] S. Bose, R. Sur Chowdhury, R. Das, and U. Maulik, "Dense Dilated Deep Multiscale Supervised U-Network for biomedical image segmentation," *Computers in Biology and Medicine*, vol. 143, p. 105274, Apr. 2022, doi: 10.1016/j.compbiomed.2022.105274.

[12] C. Cui *et al.*, "Diabetic Wound Segmentation using Convolutional Neural Networks," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Berlin, Germany: IEEE, Jul. 2019, pp. 1002–1005. doi: 10.1109/EMBC.2019.8856665.

[13] T. Lan, Z. Li, and J. Chen, "FusionSegNet: Fusing global foot features and local wound features to diagnose diabetic foot," *Computers in Biology and Medicine*, vol. 152, p. 106456, Jan. 2023, doi: 10.1016/j.compbiomed.2022.106456.

[14] D. Bouallal *et al.*, "STANDUP database of plantar foot thermal and RGB images for early ulcer detection," *Open Res Europe*, vol. 2, p. 77, Jun. 2022, doi: 10.12688/openreseurope.14706.1.

[15] S. Gupta and S. K. Pahuja, "Detection of Ischemia in DFU to acknowledge the wound status from a far-off location," in *2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)*, 2021, pp. 1–7. doi: 10.1109/SMARTGENCON51891.2021.9645778.

[16] L. Wang *et al.*, "An Automatic Assessment System of Diabetic Foot Ulcers Based on Wound Area Determination, Color Segmentation, and Healing Score Evaluation," *Journal of Diabetes Science and Technology*, vol. 10, no. 2, pp. 421–428, 2016, doi: 10.1177/1932296815599004.

[17] R. Niri, H. Douzi, Y. Lucas, and S. Treuillet, "A Superpixel-Wise Fully Convolutional Neural Network Approach for Diabetic Foot Ulcer Tissue Classification," in *Pattern Recognition. ICPR International Workshops and Challenges*, Cham: Springer International Publishing, 2021, pp. 308–320.

[18] E. Gutierrez, B. Castañeda, S. Treuillet, and I. Hernandez, "Multimodal and Multiview Wound Monitoring with Mobile Devices," *Photonics*, vol. 8, no. 10, p. 424, Oct. 2021, doi: 10.3390/photonics8100424.

[19] S. Yang *et al.*, "Sequential Change of Wound Calculated by Image Analysis Using a Color Patch Method during a Secondary Intention

Healing.," *PLoS One*, vol. 11, no. 9, p. e0163092, 2016, doi: 10.1371/journal.pone.0163092.

[20] C. Wang *et al.*, "A unified framework for automatic wound segmentation and analysis with deep convolutional neural networks," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug. 2015, pp. 2415–2418. doi: 10.1109/EMBC.2015.7318881.

[21] M. Kręcichwost *et al.*, "Chronic wounds multimodal image database," *Computerized Medical Imaging and Graphics*, vol. 88, p. 101844, Mar. 2021, doi: 10.1016/j.compmedimag.2020.101844.

[22] C. Wang *et al.*, "Fully automatic wound segmentation with deep convolutional neural networks," *Sci Rep*, vol. 10, no. 1, p. 21897, Dec. 2020, doi: 10.1038/s41598-020-78799-w.

[23] C. Wang *et al.*, "FUSeg: The Foot Ulcer Segmentation Challenge," *Information*, vol. 15, no. 3, 2024, doi: 10.3390/info15030140.

[24] B. Cassidy *et al.*, "Dfuc2020: Analysis towards diabetic foot ulcer detection," *arXiv preprint arXiv:2004.11853*, 2020.

[25] M. H. Yap *et al.*, "Deep learning in diabetic foot ulcers detection: A comprehensive evaluation," *Computers in Biology and Medicine*, vol. 135, p. 104596, Aug. 2021, doi: 10.1016/j.compbiomed.2021.104596.

[26] M. H. Yap *et al.*, "Diabetic foot ulcers segmentation challenge report: Benchmark and analysis," *Medical Image Analysis*, vol. 94, p. 103153, May 2024, doi: 10.1016/j.media.2024.103153.

[27] T.-Y. Liao, C.-H. Yang, Y.-W. Lo, K.-Y. Lai, P.-H. Shen, and Y.-L. Lin, "HarDNet-DFUS: Enhancing Backbone and Decoder of HarDNet-MSEG for Diabetic Foot Ulcer Image Segmentation," in *Diabetic Foot Ulcers Grand Challenge*, M. H. Yap, C. Kendrick, and B. Cassidy, Eds., Cham: Springer International Publishing, 2023, pp. 21–30.

[28] H. Yi *et al.*, "OCRNet for Diabetic Foot Ulcer Segmentation Combined with Edge Loss," in *Diabetic Foot Ulcers Grand Challenge: Third Challenge, DFUC 2022, Held in Conjunction with MICCAI 2022, Singapore, September 22, 2022, Proceedings*, Berlin, Heidelberg: Springer-Verlag, 2023, pp. 31–39. doi: 10.1007/978-3-031-26354-5_3.

[29] A. Galdran, G. Carneiro, and M. A. G. Ballester, "On the Optimal Combination of Cross-Entropy and Soft Dice Losses for Lesion Segmentation with Out-of-Distribution Robustness," in *Diabetic Foot Ulcers Grand Challenge*, vol. 13797, M. H. Yap, C. Kendrick, and B. Cassidy, Eds., in Lecture Notes in Computer Science, vol. 13797. , Cham: Springer International Publishing, 2023, pp. 40–51. doi: 10.1007/978-3-031-26354-5_4.

[30] Y.-H. Chen, Y.-J. Ju, and J.-D. Huang, "Capture the Devil in the Details via Partition-then-Ensemble on Higher Resolution Images," in *Diabetic Foot Ulcers Grand Challenge*, vol. 13797, M. H. Yap, C. Kendrick, and B. Cassidy, Eds., in Lecture Notes in Computer Science, vol. 13797. , Cham: Springer International Publishing, 2023, pp. 52–64. doi: 10.1007/978-3-031-26354-5_5.

[31] R. Brüngel, S. Koitka, and C. M. Friedrich, "Unconditionally Generated and Pseudo-Labeled Synthetic Images for Diabetic Foot Ulcer Segmentation Dataset Extension," in *Diabetic Foot Ulcers Grand Challenge*, vol. 13797, M. H. Yap, C. Kendrick, and B. Cassidy, Eds., in Lecture Notes in Computer Science, vol. 13797. , Cham: Springer International Publishing, 2023, pp. 65–79. doi: 10.1007/978-3-031-26354-5_6.

[32] M. S. A. Toofanee *et al.*, "DFU-SIAM a Novel Diabetic Foot Ulcer Classification With Deep Learning," *IEEE Access*, vol. 11, pp. 98315–98332, 2023, doi: 10.1109/ACCESS.2023.3312531.

[33] M. H. Yap, B. Cassidy, J. M. Pappachan, C. O'Shea, D. Gillespie, and N. D. Reeves, "Analysis Towards Classification of Infection and Ischaemia of Diabetic Foot Ulcers," in *2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*, Athens, Greece: IEEE, Jul. 2021, pp. 1–4. doi: 10.1109/BHI50953.2021.9508563.

[34] Y. P. Toledo, A. L. D. S. Pereira, A. P. Quesada, R. F. De Moraes, S. H. Garcia, and L. A. F. Fernandes, "Scalable Segmentation of Diabetic Foot Ulcers," in *2024 IEEE 37th International Symposium on Computer-Based Medical Systems (CBMS)*, Guadalajara, Mexico: IEEE, Jun. 2024, pp. 514–520. doi: 10.1109/CBMS61543.2024.00091.

[35] S. Nagaraju *et al.*, "Automated Diabetic Foot Ulcer Detection and Classification Using Deep Learning," *IEEE Access*, vol. 11, pp. 127578–127588, 2023, doi: 10.1109/ACCESS.2023.3332292.

[36] S.-J. Kuo, P.-H. Huang, C.-C. Lin, J.-L. Li, and M.-C. Chang, "Improving Limited Supervised Foot Ulcer Segmentation Using Cross-Domain Augmentation Strategies," in *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Seoul, Korea, Republic of: IEEE, Apr. 2024, pp. 2011–2015. doi: 10.1109/ICASSP48485.2024.10446498.

[37] A. Kirillov *et al.*, "Segment Anything," 2023, *arXiv*. doi: 10.48550/ARXIV.2304.02643.

[38] Y.-P. Chen, Q.-C. Long, H.-J. Wang, S.-S. Tang, and C.-Y. Lee, "A Deep Learning-Based Segmentation Strategy for Diabetic Foot Ulcers: Combining the Strengths of HarDNet-MSEG and SAM Models," in *2023 Sixth International Symposium on Computer, Consumer and Control (IS3C)*, Taichung, Taiwan: IEEE, Jun. 2023, pp. 378–381. doi: 10.1109/IS3C57901.2023.00107.

[39] L. Taipe, J. Bardales, K. Pena-Pena, G. Comina, and M. Segovia, "A Hybrid Approach Incorporating Superpixels for Diabetic Foot Lesion Segmentation Using YOLOv5 and SAM," in *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, Athens, Greece: IEEE, May 2024, pp. 1–4. doi: 10.1109/ISBI56570.2024.10635816.

[40] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds., Cham: Springer International Publishing, 2015, pp. 234–241.

[41] K. Simonyan and A. Zisserman, *Very Deep Convolutional Networks for Large-Scale Image Recognition*. 2014.

[42] M. Tan and Q. V. Le, "EfficientNetV2: Smaller Models and Faster Training," 2021, doi: 10.48550/ARXIV.2104.00298.

[43] S. Rezvani, M. Fateh, and H. Khosravi, "ABANet: Attention boundary-aware network for image segmentation," *Expert Systems*, vol. 41, no. 9, p. e13625, Sep. 2024, doi: 10.1111/exsy.13625.

[44] T. Fan, G. Wang, Y. Li, and H. Wang, "MA-Net: A Multi-Scale Attention Network for Liver and Tumor Segmentation," *IEEE Access*, vol. 8, pp. 179656–179665, 2020, doi: 10.1109/ACCESS.2020.3025372.

[45] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," 2019, doi: 10.48550/ARXIV.1905.11946.

[46] A. Chaurasia and E. Culurciello, "LinkNet: Exploiting encoder representations for efficient semantic segmentation," in *2017 IEEE Visual Communications and Image Processing (VCIP)*, St. Petersburg, FL: IEEE, Dec. 2017, pp. 1–4. doi: 10.1109/VCIP.2017.8305148.

[47] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation," 2018, *arXiv*. doi: 10.48550/ARXIV.1802.02611.

[48] B. Cassidy *et al.*, "The DFUC 2020 Dataset: Analysis Towards Diabetic Foot Ulcer Detection.," *touchREV Endocrinol*, vol. 17, no. 1, pp. 5–11, Apr. 2021, doi: 10.17925/EE.2021.17.1.5.

[49] M. Goyal, N. D. Reeves, S. Rajbhandari, and M. H. Yap, "Robust Methods for Real-Time Diabetic Foot Ulcer Detection and Localization on Mobile Devices," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 4, pp. 1730–1741, Jul. 2019, doi: 10.1109/JBHI.2018.2868656.

[50] M. Goyal, N. D. Reeves, A. K. Davison, S. Rajbhandari, J. Spragg, and M. H. Yap, "DFUNet: Convolutional Neural Networks for Diabetic Foot Ulcer Classification," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 4, no. 5, pp. 728–739, 2020, doi: 10.1109/TETCI.2018.2866254.

[51] I. C. Dipto *et al.*, "Quantifying the Effect of Image Similarity on Diabetic Foot Ulcer Classification," in *Diabetic Foot Ulcers Grand Challenge*, Cham: Springer International Publishing, 2023, pp. 1–18. doi: https://doi.org/10.1007/978-3-031-26354-5_1.

STUARDO LUCHO received the B.S. degree in Telecommunications Engineering from Pontificia Universidad Católica del Perú, in 2011 and the M.S. degree in Computer Science from Pontificia Universidad Católica del Perú, Lima, Perú, in 2017.

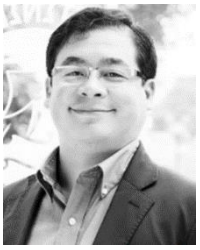
She is currently pursuing the Ph.D. degree in computer science with the PRISME Laboratory in the University of Orléans, France. His research interests include semantic segmentation in wounds and stone-by-stone segmentation using deep learning methods and computer vision techniques, cloud computing and internet of things.

ROOZBEH NAEMI is a Professor in Rehabilitation and Assistive Technology at the University of Salford. He is an accomplished engineer and scientist with a distinguished career in clinical research and an interdisciplinary educational background spanning from mechanical and biomedical engineering to biomechanics. Based at the Centre for Human Movement and Rehabilitation in Roozbeh joined the University of Salford in 2024.

A Chartered Scientist through the Institute of Physics and Engineering in Medicine (IPEM) and a Chartered Engineer through the Institute of Mechanical Engineers (IMechE), Roozbeh is actively involved in professional societies. He sits on the editorial panels of international journals and serve as an expert reviewer for national and international funding bodies.

Roozbeh's research interests encompass a wide spectrum, from soft tissue biomechanics using medical imaging to utilising AI in biomechanics. He also published in viscoelasticity imaging, computational modelling, and tissue microcirculation. Roozbeh's primary focus lies in the development of clinically viable biomechanical assessments for improving the prognosis of diabetic foot disease.

Beyond his immediate research focus, Roozbeh is keenly interested in big health data to develop predictive models. His enthusiasm extends to collaborating on projects involving healthcare, medical, and rehabilitation technologies.



BENJAMIN CASTAÑEDA (Senior Member, IEEE) received the B.S. degree in electronics engineering from the Pontificia Universidad Católica del Perú (PUCP), Lima, Peru, in 2000, the M.S. degree in computer engineering from the Rochester Institute of Technology, Rochester, NY, USA, in 2004, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Rochester, in 2006 and 2009, respectively.

He is currently the Chair of the Biomedical Engineering Program at PUCP. He has more than 15 years of experience on biomedical ultrasound and medical imaging analysis. His research interests include quantitative elastographic imaging (breast cancer diagnosis, prostate cancer detection, and skin characterization), computed aided diagnosis tools (automated diagnosis of Tuberculosis, spondyloarthrosis, follow-up of treatment for Leishmaniasis, and preventive diagnosis in maternal-perinatal health), and telemedicine (obstetric ultrasound and colposcopy).

SYLVIE TREUILLET received the master's degree in electrical engineering from the University of Clermont-Ferrand, France, in 1988, and the Ph.D. degree in computer vision from the University of Clermont-Ferrand, France, in 1993. She is a full Professor at the engineering school of University of Orléans, France. Her research interests within the PRISME Laboratory include computer vision for 3-D object modeling, deep learning, and multimodal image analysis for industrial or medical applications.