IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 18 June 2022, accepted 18 July 2022, date of publication 28 July 2022, date of current version 3 August 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3194529

SURVEY

A Survey of Wound Image Analysis Using Deep Learning: Classification, Detection, and Segmentation

RUYI ZHANG[1], DINGCHENG TIAN[1], DECHAO XU[1], WEI QIAN[1,2], AND YUDONG YAO[1], (Fellow, IEEE)

[1]Research Institute for Medical and Biological Engineering, Ningbo University, Ningbo 315211, China
[2]College of Medicine and Biological Information Engineering, Northeastern University, Shenyang 110819, China

Corresponding author: Yudong Yao (yaoyudong@nbu.edu.cn)

ABSTRACT Wounds not only harm the physical and mental health of patients, but also introduce huge medical costs. Meanwhile, there is a shortage of physicians in some areas, and clinical examinations are sometimes unreliable in wound diagnosis. Reliable wound analysis is of great importance in its diagnosis, treatment, and care. Currently, deep learning has developed rapidly in the field of computer vision and medical imaging and has become the most commonly used technique in wound image analysis. This paper studies the current research on deep learning in the field of wound image analysis, including classification, detection, and segmentation. We first review the publicly available datasets from various research, and study the preprocessing methods used in wound image analysis. Second, various models used in different deep learning tasks (classification, detection, and segmentation) and their applications in different types of wounds (e.g., burns, diabetic foot ulcers, pressure ulcers) are investigated. Finally, we discuss the challenges in the field of wound image analysis using deep learning, and provide an outlook on the research and development prospects.

INDEX TERMS Deep learning, wound image, classification, detection, segmentation.

I. INTRODUCTION

As a "silent epidemic" [1], wounds not only cause severe physical pain to individual patients, such as background pain caused by the wound itself and operative pain from clinical interventions [2], but also introduce a certain degree of psychological impact, such as worry and anxiety in patients who suffer from traumatic pain, and in severe cases, it may lead to depression [3]. In addition, since chronic wounds take a long time to heal, patients must undergo continuous care to prevent infection and the ongoing diagnosis and treatment of wounds place a significant financial burden on individuals as well as the society. In the UK, the National Health Service (NHS) spends €1.94 billion and €89.6 million annually on managing leg ulcers and burns, respectively [4]. At the same time, there is a shortage of surgeons in some regions, such as a lack of rotating backups after prolonged physician surgeries, or physicians are busy with other activities outside the hospital, resulting in patients not receiving timely, high-quality acute surgical care [5], [6]. Furthermore, it has been shown that clinical examinations are sometimes unreliable in the diagnosis of infections in chronic wounds, even with the participation of experienced physicians [7], as well as in acute wounds [8]. Therefore, there is a need for a low-cost, rapid, and accurate wound assessment technique, such as medical imaging based methods, to provide assistance in wound diagnosis, prognosis, care, and other related tasks.

Along with the rapid development of smartphones, computer hardware, and Internet techniques, research on wound or wound image assessment has started to emerge, including real-time monitoring [9], remote diagnosis [10], and mobile care [11]. Wound images can provide valuable information for an expert to accurately diagnose wounds. However, manual evaluation through wound images is time-consuming

The associate editor coordinating the review of this manuscript and approving it for publication was Kumaradevan Punithakumar.

and usually requires a significant amount of experience [12] and training an experienced physician is costly in terms of time. Researchers have made great efforts to address this issue and various solutions have been proposed to assist physicians in wound diagnosis through wound images. Traditional digital image processing using machine learning is one of the most commonly used techniques for wound diagnosis [13], [14], but it has high time costs, since when describing the characteristics of different target images, a large number of parameters need to be manually adjusted, such as using line search techniques to tune free parameters in the support vector machine (SVM) that control the penalty of the classification error [15], and grid search techniques to select the combination of network size and weight decay that give the feed-forward neural networks (NN) the best performance [16]. With a sufficient amount of labeled training data, deep learning techniques can now effectively address this problem. The potential of deep learning in image processing has been widely recognized since the AlexNet architecture based on convolutional neural networks (CNN) achieves impressive results in the ImageNet competition. The CNN model is the most commonly used model in deep learning [17]. It has the advantage that it can automatically extract multiple levels of image visual features, and does not need to manually adjust a large number of parameters [18], which effectively improves the efficiency of image processing tasks. With the availability of more and more publicly available datasets, deep learning has made rapid progress in the field of medical imaging [19], including the wound image analysis [20], and diagnostic tools based on deep learning frameworks have proven to be effective in aiding clinical decision-making [21].

Zahia *et al.* [22] publish a review of machine learning techniques in pressure injury in 2019. Anisuzzaman *et al.* [23] publish a review of artificial intelligence techniques in wound assessment in 2021, including a review of rule-based algorithms, machine learning algorithms, and deep learning algorithms. Although these studies cover a large amount of work, we believe that the review in the field of deep learning is not comprehensive. Different from previous work, we provide a more comprehensive overview of deep learning methods, including a review of publicly available datasets used in deep learning tasks, an introduction for data preprocessing methods, and various deep learning models. At the same time, we review the latest research in the field of deep learning as applied to various types of wounds.

We retrieve more than 90 research papers through Google scholar searches using the query terms "deep learning", "classification", "detection", "segmentation", "wound" and combinations of various disease names. After determining specific wound types and deep learning tasks, we carefully screen a total of approximately 50 papers considering the publication date and the number of citations. 64% of the papers were published after 2020.

This paper focuses on a comprehensive review of the applications of deep learning in wound image analysis, including classification, detection, and segmentation. We review the available wound image datasets and study the data preprocessing methods for the wound image analysis. We examine several of the most common deep learning models for the wound image analysis and introduce some model improvements for the analysis of wound images. Table 1 describes the coverage of this wound image research survey paper, including deep learning tasks, wound analysis applications, and wound types. We show in Fig. 1 some different types of wound images from the Medetec dataset [24].

This paper is organized as follows. In Section II, public wound image datasets are introduced, as well as methods for data preprocessing. In Section III, the commonly used deep learning model architectures and the model modification for wound and evaluation metrics are introduced. In Section IV-VI, specific wound analysis applications in various wound types are presented. Discussions and conclusions are given in Section VII and VIII.

TABLE 1. **Wound image research using deep learning.**

Deep learning task	Wound analysis applications	Wound types
Classification	Identification Severity classification Tissue classification	Burn Diabetic foot ulcer Diabetic wound Other types of wounds
Detection	Wound detection	Diabetic foot ulcer Skin wound Chronic wound
Segmentation	Region segmentation Depth segmentation Tissue segmentation	Burn Diabetic foot ulcer Pressure ulcer Other types of wounds

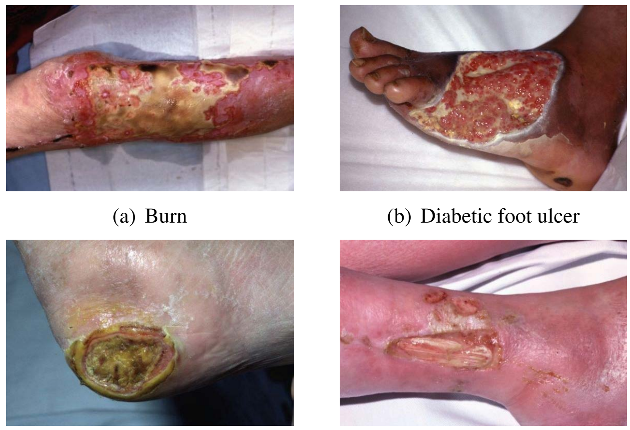
(a) Burn (b) Diabetic foot ulcer (c) Pressure ulcer (d) Venous ulcer

FIGURE 1. **Different types of wound images.**

II. DATA AND DATA PREPROCESSING

A. DATA

Data is one of the most important parts of deep learning. When deep learning models are trained with many

TABLE 2. Wound datasets.

Dataset	Wound type	Size	Labeled	Paper
Medetec dataset [24]	Mixed	-	No	[25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35]
BIP_US database [36]	Burn	94	Yes	[37], [38]
FUSeg dataset [28]	Diabetic foot ulcer	1210	Yes	[29]
Sårwebben [39]	Mixed	-	No	[40]
Chronic wound database [41]	Chronic wound	188	Yes	[42]
AHZ dataset [28]	Diabetic foot ulcer	1109	Yes	[28]
AHZ&UWM dataset [34]	Mixed	538	Yes	[34]

parameters, but the amount of data used for training is insufficient, the network model is prone to overfitting [43]. In the field of medical images, collecting datasets is challenging. First, medical images are more difficult to share publicly due to privacy protection. With the continuous development of medical data analysis methods, even after anonymizing the images, hackers may still be able to identify patients through technical means. Therefore, data in most studies are not publicly available [44]. Second, the primary job of medical professionals is not data collection, and the acquisition of a batch of images may be done by multiple personnel, which can lead to inconsistent standards of the collected images. In addition, due to the different imaging equipment, distance and angle of the capture, the image content can show significant differences, including color mode, light, intensity, edges, etc., making the network model need more parameters to analyze the images [45]. Finally, some medical images cannot be acquired in large quantities, since the capture method can harm the patient's body [46]. Producing datasets is also a very difficult task. Annotating images usually needs to be done manually and, even if the task is performed by people with extensive professional experience, the standards are not necessarily identical across experts. In addition, data imbalance is a common problem in medical images, where the proportion of negative samples in the dataset may be much larger than positive samples, which can lead to bias in the training of the model [47].

Public datasets have been very helpful in advancing the research of deep learning in wound images. First, publicly available datasets of the same wound type can be used directly for data augmentation to improve model performance. Second, Ohura *et al.* [48] train a deep learning model using a dataset of pressure ulcers (PU), and test it on diabetic foot ulcers (DFU) and venous leg ulcers (VLU) data. The results demonstrate that datasets with different wound types can also be used for model training. In addition, common datasets allow for fair comparisons among different deep learning systems. Some of the publicly available datasets are summarized in Table 2.

B. DATA PREPROCESSING

Although deep learning models can be trained directly based on original images when the data is sufficiently clear and of low noise [17], the training performance of the model still varies depending on data preprocessing methods [53]. Data shortage is one of the common problems in deep learning in the field of wound applications. This is due to the lack of public datasets for some wound types and the difficulty of obtaining the sufficient amount of data through medical institutions. Data augmentation is widely used in preprocessing as a method to expand the number of samples without substantially increasing the existing data. Conventional image augmentation methods include geometric transformation, i.e., rotation, flipping, random scaling, etc., and color transformation, i.e., contrast transformation, color model conversion, Gaussian blur, etc.

Image generation is an effective way to expand the training set. Zhang *et al.* [35] propose an image generation system based on deep convolutional generative adversarial networks (DCGANs) that can generate simulated chronic wound images. The system consists of a generator architecture and a discriminator architecture, and the model is trained using mini-batch stochastic gradient descent. Experimental comparison with the original data demonstrates that the images generated through this system can effectively improve the model segmentation accuracy. Dai *et al.* [52] propose a framework capable of automatically generating burn images with annotations. The burn wounds are first generated using a style-generative adversarial network (Style-GAN), which is trained based on the non-saturating loss with regularization function. Then the wounds are fused with human skins using color adjusted seamless cloning (CASC), and finally, the burn annotated dataset is obtained using coordinate transformation in 3D space.

Light reflections generated during shooting are one of the reasons for the degradation of wound image quality, and Rajathi *et al.* [49] use a flash light removal method to eliminate such interference. Firstly, the original image is converted to a grayscale image based on linear combination and non-wounded regions are identified. After converting the grayscale image to a binary image, white pixel areas are detected and filled through a given threshold. In contrast, Wagh *et al.* [50] improve the stability of the algorithm by using the contrast limited adaptive histogram equalization (CLAHE) method to enhance the interference of light and noise in the image. The method requires the user to define

TABLE 3. Data preprocessing.

Approach	Method	Paper
Geometric transformation	Rotation, flipping, cropping, random scaling, etc.	-
Color transformation	Contrast transformation, normalization, sharpening, color model conversion, Gaussian blur, etc.	-
	Flash light removal	[49]
	Contrast limited adaptive histogram equalization	[50]
	Color enhancement	[32]
	Color space reduction	[51]
Image generation	DCGANs	[35]
	Style-GAN	[52]

a limit to be used as the maximum allowable local contrast enhancement factor. Within the clip range, the noise of the image is enhanced appropriately.

Another challenge in wound image processing is the color variation caused by the external environment. Pholberdee *et al.* [32] use a color enhancement method. This method obtains color mapping parameters between cameras through a deep learning model and generates images based on these parameters to extend the dataset. Since RGB space is not sufficient for comparison between colors in 3D space, Godeiro *et al.* [51] use a color space reduction approach. The RGB components are first quantized, and the histogram colors are reduced from 2563 to 643, and then the components are converted to a CIELab color space, thus allowing direct color comparisons in 3D color space based on geometric separability.

III. DEEP LEARNING METHODS

A. CNN, VGG, AND RELATED MODELS

CNN has been widely used due to its excellent performance and efficiency in image processing. In 1989, LeCun *et al.* [54] propose CNN for handwritten character recognition. AlexNet [55] is the first modern deep convolutional neural network (DCNN) model, which first applied techniques such as ReLU, Dropout and GPU operation acceleration in CNN, and achieved excellent performance. With AlexNet winning the ImageNet competition in 2012 with a far superior first place, CNN is able to rapidly spread to various application fields. VGG [56] adopts a larger number of small convolutional kernels in the convolutional layer instead of the otherwise larger ones, thus reducing the number of parameters as well as increasing the number of nonlinear mappings, and significantly improve the classification performance of the network. DeepLab [57] replaces the ordinary convolution of VGG with atrous convolution for segmentation tasks, and then performs post-processing optimization on the obtained segmentation results through Conditional Random Field (CRF). MobileNet [58] uses depthwise separable convolution to deepen the network. Compared with the VGG-16 network, MobileNet can better reduce the parameters and calculation amount while slightly reducing the model accuracy. Compared to previous work, the residual network (ResNet) [59] is notable for its network depth. ResNet addresses the degradation problem of deep networks by introducing a residual unit through a shortcut connection. Compared with VGG, ResNet has a great improvement in computational speed and model accuracy. DenseNet [60] establishes a dense connection between all the previous layers and the latter layer, and realizes feature reuse through the connection of features on the channel. These features allow DenseNet to achieve better performance than ResNet with fewer parameters and computational costs.

Bhansali and Kumar [37] design an 8-layer CNN network for the classification of burns, which is named BurnNet. In the preprocessing stage, they use an anti-aliasing technique to adjust the image size and perform data augmentation through affine transformation. At the end of the network, two connected dense layers enhance the propagation of features and reduce the number of parameters to be calculated. Ong *et al.* [61] propose a CNN-based encoder-decoder architecture to segment chronic wounds. They use more down-sampling and up-sampling layers to form an encoder block and a decoder block, respectively, and mapped the features of the encoder. The up-sampled feature maps are concatenated to form a ladder-like architecture that enables the entire model to be trained end-to-end. Compared with general models, this structure has higher accuracy and faster computation speed in pixel-level wound segmentation.

B. R-CNN, YOLO, AND RELATED MODELS

The two-stage detection algorithm divides the detection problem into two stages. First, regions of interest (RoIs) are extracted from the whole image, and then each RoI is updated as well as classified. The R-CNN [62] is one of the first deep learning frameworks to use the region proposal-based approach. Since R-CNN takes longer time in extracting feature vectors, Fast R-CNN [63] circumvents the redundant feature extraction operations in R-CNN. Fast R-CNN uses the RoI pooling layer instead of the last max pooling layer and performs only one global feature extraction for the whole image. Faster R-CNN [64] integrates the region proposal network (RPN) with Fast R-CNN, so that region detection, feature extraction, and window classification are all done in one network, which greatly improves the target detection speed.

The one-stage detection algorithm directly generates target locations and class probabilities, which improves the computational speed at the expense of accuracy compared to the two-stage detection algorithm. YOLO [65] draws on the GoogLeNet network structure to implement end-to-end object detection using a separate CNN model. Compared with Faster R-CNN, YOLO produces fewer false positives in the background region [66]. Single shot multi-box detector (SSD) [67] uses VGG-16 as the base model and adds more convolutional layers to the network. Compared with YOLO, SSD uses multiscale feature maps for detection, which has the advantage of allowing relatively large feature maps to be used for detecting relatively small targets, and can make up for the shortcomings of YOLO in small object detection. Different from the previous detection framework, Detection Transformer (DETR) [68] adopts the transformer as the core architecture, and regards target detection as an ensemble prediction problem, which shows better performance in the detection task of large objects.

Oliveira *et al.* [69] improve the model on the basis of Faster-RCNN. By adjusting the detector parameters, the range of the training region is expanded. The number of anchors is also increased, allowing the network to detect smaller lesions. In the DFU detection task, faster detection speed and better detection accuracy are obtained. Amin *et al.* [70] combine ShuffleNet with the YOLOv2 model for more efficient wound detection. In the localization stage, the images are first fed into ShuffleNet with 172 layers, and then passed to the YOLOv2 model for detection through the ReLU node-199 layer. At the same time, they design a model with a 16-layer CNN capable of classifying DFU images into two conditions, ischemia and infection.

C. FCN AND RELATED MODELS

Fully convolutional networks (FCN) [76] is a milestone in the field of deep learning in image segmentation. FCN does not contain fully connected layers, and the model consists of only convolutional layers. The advantage of this structure is that it can adapt to images of any scale as input. At the same time, a skip structure is used to combine the results of the last layer with the results of the shallow layer. This effectively addresses the problem of loss of details in the early stage. Compared to FCN, U-Net [77] also follows the structure of an encoder-decoder, but uses multi-level convolution on the decoder and uses a concatenation operation in the skip connection, whereas FCN uses a summation operation. U-Net can be efficiently trained with a small number of samples with data augmentation, which is very effective in biomedical segmentation tasks. Mask R-CNN [78] adds a predictive segmentation mask to the Faster R-CNN. In the network, Mask R-CNN replaces the RoI pooling layer in Faster R-CNN with RoI align layer, so that smaller feature maps can be output to make the output pixels correspond precisely to the input pixels. The Mask R-CNN can perform both target detection and segmentation with faster computation speed and better robustness than the FCN method.

Pathompatai *et al.* [31] design a region-focused training strategy in U-Net model training to overcome the shortcomings of deep learning models in processing large images. The strategy first divides the image into several sub-images, and defines the region where the wound is located as the central neighborhood. In the central neighborhood, the number of cropped sub-images will be more than other regions, so that the model will be trained more times in the wound region, thereby improving the overall training accuracy of the model, and the reliability of the strategy is verified by comparing the training results of the model with U-Net. Jiao *et al.* [79] propose a Mask R-CNN framework with Residual Network-101 as the backbone. In the framework, atrous convolution is used to expand the field of view of the model, making the model more suitable for datasets with different burn depths and wound sizes, and the number of feature maps is increased in the RPN network, whereas the loss function of class branching is changed.

D. EVALUATION METRICS

The performance of deep learning models can be evaluated through standard metrics. Using different evaluation metrics can make more comprehensive comparisons between models and provide researchers with appropriate directions for optimizing models [80]. The evaluation metrics can be defined by confusion matrices, where the positive and negative instances of correct prediction are denoted as True positive (TP) and True negative (TN), and the negative and positive instances of incorrect prediction are denoted as False positive (FP) and False negative (FN), respectively.

Accuracy represents the percentage of correctly classified samples among the total samples, and is the most basic metric for evaluating model performance. The performance of the model can be accurately measured in the class balanced cases. But in the case of class imbalance, using accuracy creates performance evaluation limitations.

$$\mathrm{Accuracy} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN}} \tag{1}$$

Precision represents the proportion of correctly classified positive samples in the prediction results among the actual positive samples, and this metric is sensitive to over-prediction.

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{2}$$

The recall represents the proportion of correctly classified positive samples in the prediction results in the overall predicted positive samples, and this metric is sensitive to under-prediction.

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{TN}} \tag{3}$$

TABLE 4. Deep learning in classification of burns.

Application	Target class	Dataset size	DL model	Performance (%)				Paper
				Accuracy	Precision	F1-Score	Recall	
Severity classification	Three types of burn degree	450	DCNN	79.4	-	-	-	[71]
	Three types of burn degree	5637	ResNet50	-	> 81	> 78	> 73	[72]
	Superficial dermal, deep dermal, and full thickness	94	AlexNet	77.8	-	-	-	[38]
	Full-thickness, deep dermal, and superficial dermal	94	BurnNet	99.87	-	99.87	-	[37]
	Low, moderate, and severe	4614	ResNet50	84.85	78.0	77.8	77.6	[73]
	Four types of burn depth, normal skin, and background	23	ResNet101	90.54	75.19	-	-	[74]
Identification	Burn, and healthy	1900	ResNet50	> 97.1	-	-	-	[75]
	Graft, and nonGraft	94	AlexNet	90.5	90.6	89.22	87.9	[38]

F1-Score is calculated by combining precision and recall, which can reflect the degree of boundary matching between the predicted result and the true value. The F1-Score is also known as the Dice Similarity Coefficient (DSC).

$$\text{F1-Score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

$$\text{DSC} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \tag{5}$$

Sensitivity represents the proportion of correct predictions in actual positive samples, which can reflect the predictive performance of the model for positive samples.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{6}$$

Specificity represents the proportion of correct predictions in actual negative samples, which can reflect the prediction performance of the model for negative samples.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{7}$$

Intersection over Union (IoU) can reflect the correlation between actual results and predicted results.

$$\text{IoU} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \tag{8}$$

Mean average precision (mAP) is a commonly used evaluation metric in object detection tasks. First, the average precision of each class is calculated based on precision and recall, and then the average precision of all classes is calculated.

In addition, in a multi-class prediction model, the above evaluation metrics are usually generalized to the average of the metrics of each class.

IV. WOUND CLASSIFICATION

A. CLASSIFICATION OF BURNS

Burns are a very serious type of wounds, and every year a large number of people are disabled or even die due to burns. Severity assessment of wounds is an initial preparation for burn wound diagnosis, monitoring, and care sessions and is the most common burn injury classification task. Chauhan and Goyal [73] propose a deep learning framework that can classify images into four body parts, and three burn levels. It first trains the M-ResNet50 model using non-burn images to predict body parts, and then feeds wound images of specific body parts to train the ResNet50 model to predict severity, and the framework has excellent performance in both classification tasks. Further, Cirillo *et al.* [74] use a pretrained CNN model to classify burn depth into four classes. They examine the classification results of several networks, including VGG-16, GoogleNet, ResNet-50, and ResNet-101, and compare them with clinical diagnosis results, demonstrating the reliability of deep learning in burn depth prediction work.

Burn identification is also a related topic and Abubakar *et al.* [75] perform burn wound identification based on the ResNet50 model. They replace the last layer of the pre-trained ConvNet model, freeze the lower layer of ResNet50, and use a fully connected layer in the top layer so that the whole model changes from the pre-trained 1000 output classes to 2 classes. The model achieves satisfactory results in the recognition of burn images of different ethnic subjects. Table 4 summarizes the various classification applications of deep learning models in burn wounds.

B. CLASSIFICATION OF DFUs/DIABETIC WOUNDS

Diabetes can lead to many complications, including DFUs, which are a long-term concern for patients' physical and mental health. Automated DFU severity assessment through deep learning can provide reliable decision-making recommendations, as well as time and cost savings for healthcare. Gamage *et al.* [82] propose a Wagner grading system for diabetic foot ulcers. A pre-trained CNN model is first fine-tuned by replacing the output layer and initializing the input layer weights. The model connects a global average pooling (GAP) as a feature extractor, and the extracted feature vectors are

TABLE 5. Deep learning in classification of DFUs/diabetic wounds.

Application	Target class	Dataset size	DL model	Performance (%)			Paper
				Accuracy	Precision	F1-Score	
Severity classification	Five grades for wound depth, and granulation tissue amount	1639	Bi-CNN	84.6 84.6	-	84.89 83.82	[81]
	Six grades based on wanger score	2400	DenseNet201	96.22	-	96.10	[82]
Identification	Normal, and abnormal	745	DFU_QUTNet	-	95.4	94.5	[83]
	Normal, and abnormal	397	DFU_SPNet	96.4	-	95.4	[84]
	Normal, and abnormal	397	DFUNet	92.5 ± 2.9	94.5 ± 3.2	93.9 ± 2.4	[85]
	Ischaemia, and non-ischaemia Infection, and non-infection	1459	CNN	90.3 ± 1.2 72.7 ± 2.5	91.8 ± 1.9 73.5 ± 3.6	90.2 ± 1.4 72.2 ± 2.8	[86]

TABLE 6. Deep learning in classification of other types of wounds.

Application	Target class	Dataset size	DL model	Performance (%)			Paper
				Accuracy	Sensitivity	Specificity	
Identification	Venous, and non-venous leg ulcers	2295 (Venous leg ulcers)	VGG-19	85	-	-	[40]
	Surgical, venous, and diabetic	400	DCNN	87.7	-	-	[34]
Tissue classification	Granulation, necrotic, and slough	30 (Chronic wounds)	U-Net	96.1 ± 4.1	91.28 ± 7.4	98.76 ± 2.3	[51]
	Wound mask, granulation, fibrin, and necrotic tissues	217 (Dermatological wounds)	ResNet	-	97.0 ± 0.4	97.4 ± 0.7	[87]

fed into a single hidden layer artificial neural network (ANN) classifier to predict the classification results. The system can effectively improve classification accuracy and computation time. For diabetic wounds, Zhao *et al.* [81] classify the diabetic wound depth based on the Bilinear CNN model whereas evaluating the granulation tissue. For model training, they adopt two steps, transfer learning and model fine-tuning. In the first step, all parameters of the convolutional block are frozen after transfer learning, and the softmax layer is adjusted based on the number of classifications for training. In the second step, the entire model is fine-tuned using back propagation, and finally good performance is achieved in the final classification task.

In addition to wound classification, wound identification as well as ischemia and infection identification of wounds are also applied in DFUs. Goyal *et al.* [85] propose the DFUNet structure, which combines depth and parallel convolutional layer, is able to increase the amount of image features extracted and reduce the depth of the network. Also, the number of neurons in the fully connected (FC) layer is reduced based on the two-classification task to obtain faster processing. Alzubaidi *et al.* [83] design the DFU_QUTNet structure to identify wounds. This structure increases the width of the network whereas maintaining the depth of the network, and trains the SVM classifier with the image features extracted using the model. Its classification results are compared with the state-of-the-art CNN network. Goyal *et al.* [86] perform ischemia and infection identification of wounds based on the Ensemble CNN model, and achieve a higher level of accuracy compared to traditional machine learning algorithms. Table 5 summarizes the applications of deep learning models for the classification of DFUs/Diabetic wounds.

C. CLASSIFICATION OF OTHER TYPES OF WOUNDS

The traditional method used by physicians to evaluate wounds is to analyze the tissue condition through visual inspection. Due to the wide variety of wound tissues and their complex appearance, the results obtained are often highly variable. Godeiro *et al.* [51] propose a system capable of segmenting and classifying chronic wound tissues. The preprocessed image is first segmented based on the watershed algorithm for the wound region, then the segmented image is converted to the CIELab color space, and finally, the converted tissue image is fed into the U-Net model to complete the classification. The U-Net model is initialized using a pre-trained VGG-16 model. Blanco *et al.* [87] use a divide-and-conquer strategy to overcome the problem of small wound image data, which requires the CNN model to process the superpixels instead of the entire image. The original wound

TABLE 7. Deep learning in detection of wounds.

Application	Dataset size	DL model	Performance (%)		Detection speed (ms)	Paper
			mAP	F1-Score		
Diabetic foot ulcers	7887	YOLOv2	95 (Ischemia), 90 (Infection)	-	-	[70]
	2000	Faster R-CNN	91.4	94.8	332	[69]
	4200	YOLOv5	67.52 (Base), 67.54 (Self-trained)	73.02, 74.07	-	[88]
		DETR	72.84 (Base), 71.25 (Self-trained)	73.55, 73.84		
	2535	YOLOv3	91.95	-	31	[89]
	1175	Faster R-CNN	91.8	-	-	[90]
Skin wounds	457	SSDLite MobileNetV2	86.46	-	23 (Desktop) 119 (smart phone)	[9]
Chronic wounds	188	YOLOv4	83	-	-	[42]

is first segmented using a superpixel construction algorithm, followed by feature extraction and tissue classification using a coupled CNN model.

In the application of wound identification, Nilsson *et al.* [40] use a pre-trained VGG19 model to classify venous and non-venous leg ulcers. Rostami *et al.* [34] propose an end-to-end ensemble deep learning model for wound image type classification. The model uses two classification strategies, patch-wise and image-wise. The input wound images are first combined by the outputs of the two classifiers, and then the combined results are fed to a multilayer perceptron (MLP) classifier for the final classification, which classifies the images into three wound types. Table 6 summarizes the applications of deep learning for classification of other types of wounds.

V. WOUND DETECTION

Regular monitoring of foot ulcers is an important part of diabetic foot care. Currently, this task still requires the physician to compare the initial as well as the available photographs in order to determine the stage of development of the foot ulcer. Automated detection of DFUs based on deep learning methods allows the monitoring task to be performed without clinical intervention [91]. Brüngel and Friedrich [88] compare the DFU detection performance of YOLOv5 and DETR. Also, they compare the test results using the base model and the self-training model at different confidence levels. The performance impact due to the self-training model and the potential for further optimization of the model are also discussed.

Single-target DFU detection can only locate the wound. If a more effective assessment of the stage of DFU development is desired, further discrimination of the wound status is required. Han *et al.* [89] propose a multi-target detection task for the healing status of DFUs based on the Wagner classification of diabetic feet, and classify the target regions into six different grades. They use data augmentation through the visual coherent image mixup method, and also use cosine learning rate to train the model for the training characteristics of the YOLOv3 model and obtain better average accuracy.

In addition to DFUs, Faria *et al.* [9] design a mobile application for real-time skin wound monitoring. In the wound detection part, they merge the SSDLite and MobileNetV2 architectures to make the model more lightweight. Meanwhile, the generalization ability of the model is improved using transfer learning, which demonstrates the reliability of real-time detection tasks on mobile devices. Monroy *et al.* [42] use the YOLOv4 model to detect chronic wounds for subsequent wound segmentation. Table 7 summarizes the applications of deep learning in wound detection.

VI. WOUND SEGMENTATION

A. SEGMENTATION OF BURNS

Total burn surface area (TBSA) is an important metric for assessing burns, and accurate segmentation of the wound region can be of significant help for TBSA estimation. Liu *et al.* [94] propose an end-to-end deep learning framework including two networks. The encoder network extracts semantic feature maps based on downsampling, and the decoder network fuses semantic information based on upsampling. Multiple backbone networks are trained using fusion loss function to complete the segmentation of the burn area. Meanwhile, they extend the output structure of the network to realize the calculation of TBSA. Chauhan *et al.* [92] perform segmentation of burns in DCNN. They use the concept of atrous spatial pyramid pooling (ASPP) in the pre-trained ResNet-101 model, and sample the feature maps generated using the pre-trained model using atrous convolution. This has the advantage of obtaining better performance in intensive prediction tasks.

Currently, the calculation of burn depth still requires specialized instruments for evaluation, which relies on

TABLE 8. **Deep learning in segmentation of burns.**

Application	Dataset size	DL model	Performance (%)				Paper
			Accuracy	IoU	DSC	Precision	
Region segmentation	449	DCNN	93.4	-	81.42	81.95	[92]
	929	FCN	85	67	-	-	[93]
	1150	Mask R-CNN	-	-	84.51	-	[79]
	516	FCN	94.59	84.67	91.70	-	[94]
	611	CNN	-	-	-	75.91	[95]
Depth segmentation	929	FCN	60	37	-	-	[93]
	516	FCN	66.84	51.44	67.82	-	[94]
	1800	Mask R-CNN	86.63	-	86.9	85	[96]
	100	U-Net	91.89 ± 0.6	-	91.88 ± 0.6	-	[97]

TABLE 9. **Deep learning in segmentation of DFUs.**

Application	Dataset size	DL model	Performance (%)				Paper
			Accuracy	IoU	DSC	Precision	
Region segmentation	445	U-Net	-	76.1	84.5	83	[98]
	1210	LinkNet + U-Net	-	85.51	92.07	92.68	[29]
	1176	Mask R-CNN	> 93.90	-	> 94.04	> 92.17	[99]
	92	U-Net	94.96	94.86	97.25	-	[100]
	188	U-Net	-	-	88.8 ± 1	90.2 ± 1	[42]
	1109	MobileNetv2	-	-	90.47	91.01	[28]
Tissue segmentation	219	FCN-32	92.68	-	75.74	78.07	[101]

experienced operators for implementation. Pabitha and Vanathi [96] design a neural network integrating Mask R-CNN with dense pose estimation for automated segmentation of wound depth. The network introduces a RoI-pose align module that allows the correction of human poses in images to a uniform format based on spatial correspondence. In the module, pose estimation and instance segmentation are performed simultaneously. After the segmentation of the burn region, the wounds are classified into three classes of depth using a classifier. Cirillo *et al.* [97] use a modified U-Net network with residuals for segmentation of four wound depths. They extend the dataset by using image rotation and elastic deformation techniques and achieve reliable segmentation results. Table 8 summarizes the applications of deep learning to segmentation in burns.

B. SEGMENTATION OF DFUs/DIABETIC WOUNDS

Accurate segmentation of DFUs can be useful in the task of early wound prevention as well as healing degree assessment. A simplified model is designed by Cui *et al.* [98]. The noise generated through light reflection is first reduced in the preprocessing stage, then the wound is segmented using a CNN model, and finally the wound region is refined in the post-processing stage. Monroy *et al.* [42] propose a two-step learning framework for wound detection and segmentation. The first step detects DFUs using the YOLOv4 model, and the second step connects the U-Net model to segment the wounds. The segmentation results of this framework are closer to the manual performance level than the direct application of U-Net for segmentation. Wang *et al.* [28] use a more lightweight MobileNetv2 model to segment DFUs. In the post-processing step, the segmentation results are binarized based on thresholds, whereas the connected component labeling (CCL) is used to fill the holes in the negative and positive parts, thus effectively improving the accuracy of the segmentation.

Automated wound tissue analysis can help avoid contact with wounds and reduce the risk of infection. Niri *et al.* [101] develop a smartphone-based ulcer tissue segmentation system. RoI segmentation is first performed using U-Net, followed by superpixel extraction of ROI using the simple linear iterative clustering (SLIC) algorithm. Finally, the pixels are classified on FCN-Net to obtain the final tissue segmentation results. The experimental performance shows that the system can effectively detect ulcers. Table 9 summarizes the applications of deep learning for segmentation in DFUs.

C. SEGMENTATION OF PRESSURE ULCERS

Pressure injuries, or pressure ulcers, resulting from the long-term pressure on local tissues cause tissue ulceration

TABLE 10. Deep learning in segmentation of pressure ulcers.

Application	Dataset size	DL model	Performance (%)				Paper
			Accuracy	IoU	DSC	Precision	
Region segmentation	1474	U-Net	99.0	99.9	93.4	-	[25]
	210	Mask RCNN	-	-	83	87	[33]
	2893	DeeplabV3	99.25	97.82	98.87	98.88	[102]
	440	U-Net	97.82 (Test on DFUs)	-	85	-	[48]
Tissue segmentation	193	CNN	-	-	91 ± 7	-	[26]

TABLE 11. Deep learning in classification of other types of wounds.

Application	Dataset size	DL model	Performance (%)				Paper
			Accuracy	IoU	DSC	Precision	
Region segmentation	446 (Skin ulcer)	ASURA	-	-	90 ± 13	92 ± 12	[103]
	400 (Neuropathic ulcer)	Mask-RCNN	86.32	-	-	-	[104]
	950	MobileNet	-	86.47	-	95.03	[27]
	583 (Chronic wound)	CNN	94.1	86.9	-	-	[61]
	180	U-Net	-	78.16	-	-	[31]
Tissue segmentation	240 (Chronic wound)	CNN	-	52.82	-	-	[32]
	1442 (Chronic wound)	DeepLabV3	-	-	87.6	-	[50]
	350 (Chronic wound)	AlexNet	86.4	-	-	-	[105]

and necrosis. Due to the difficulty of wound healing, the rehabilitation and care of pressure ulcers is a common problem plaguing the world. Chae *et al.* [25] add the attention module to the Residual U-Net. In the network structure, the Squeeze-Excitation (SE) block acquires the channel information and calculates the importance of each channel, and the attention block extracts the spatial features of the channel through the convolutional layer. Due to the lack of data volume, they also pre-train the model using other approximate wound images. Zahia *et al.* [33] propose a framework for segmentation and measurement of pressure injuries. The wound image segmentation is first performed in Mask R-CNN, followed by the generation of the top view of the image and the corresponding matrix of face indices in the 3D network. Finally, the segmentation results are matched with the top view. Based on the structure of a sensor, the wound segmentation results in the 3D mesh are obtained, as well as the actual measurement parameters of the wound. For deep learning for PUs diagnosis, Chang *et al.* [102] produce two datasets. The first one outputs the results of wound and re-epithelialization (re-ep) region segmentation directly using deep learning models. The second dataset is first based on the SLIC algorithm for superpixel extraction, followed by tissue classification through deep learning models. Five deep learning models are used for testing in both segmentation and classification tasks, and DeeplabV3 obtain the best results for all of them.

Accurate tissue segmentation can provide effective assistance in the diagnosis of PU. García-Zapirain *et al.* [26] propose a system for detecting and segmenting PU tissues. It includes two stages of RoI extraction and tissue segmentation, each of which uses a 3D CNN model. The wound image is first segmented externally to remove the background regions from the image. The results are then fed into the next stage to output the segmentation results for different tissues. Tests show that the system is able to reliably segment images of ulcers. Table 10 summarizes the various types of segmentation applications of deep learning in pressure ulcer wounds.

D. SEGMENTATION OF OTHER TYPES OF WOUNDS

In a wound image, the pixel color will be a gradient as the wound area extends into a healthy area. Based on this feature, Li *et al.* [27] enhance the location information of the image in a deep neural network. The location map is first obtained through a location encoder, and then input into the DNN together with the original image. Next, the model is trained using the location-enhanced convolutional kernel, and the model output is fused with the location map to obtain the final wound segmentation result. Chino *et al.* [103] design the automatic skin ulcer region assessment (ASURA) framework for segmentation and measurement of skin wounds. In the segmentation part, ASURA uses an encoder-decoder structural model that approximates the U-Net, and a heaviside

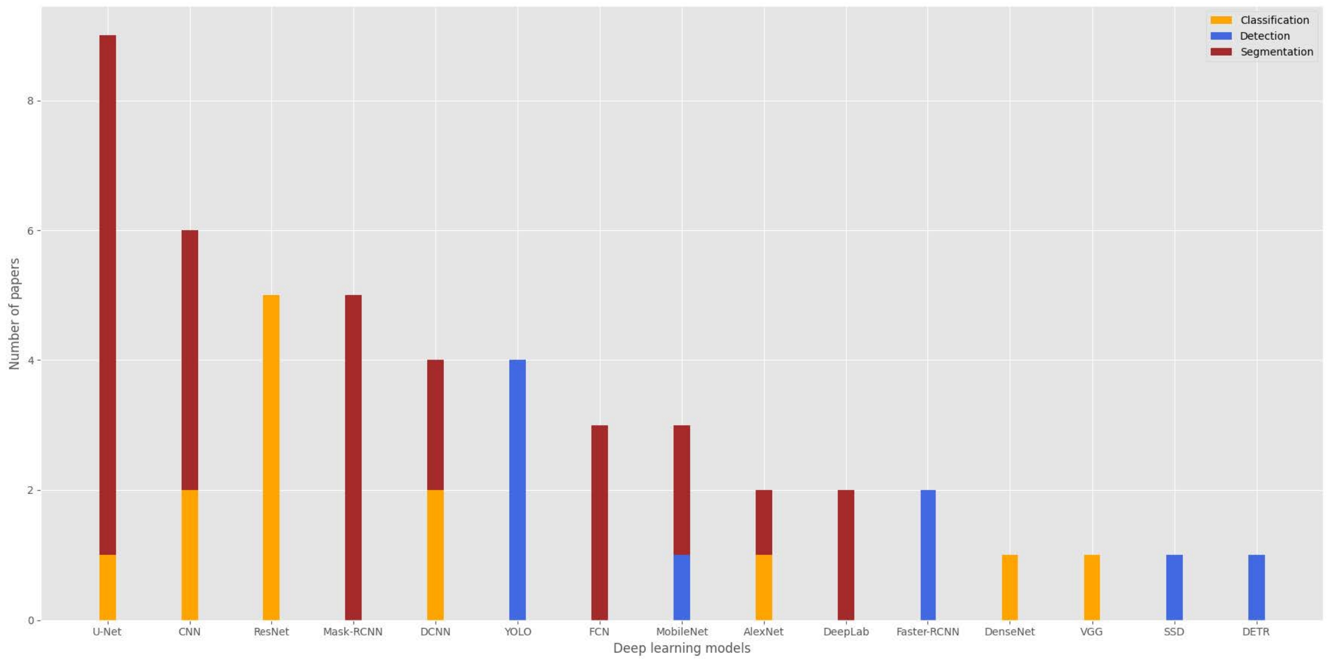

FIGURE 2. **General deep learning models used in different tasks.**

step function is used at the output to adjust the mask. In the measurement part, ASURA obtains the pixel density from a reference in the image to estimate the area of the wound.

Mobile devices can further shorten the time of nursing decision-making and wound analysis on smartphone images becomes more meaningful. Wagh *et al.* [50] use different CNN models to segment different types of wound tissue images, including FCN, U-Net, and DeepLabV3. In post-processing, they use CRF with Gaussian edge potentials to improve the continuity of the segmentation results, and the comparison results with the associated hierarchical random field (AHRF) method show that deep learning performs better in the segmentation task. In the tissue segmentation task, in order to be able to further improve the effectiveness of the decision-making system, Nejati *et al.* [105] train a supervised deep learning network to classify chronic wound tissue into 7 categories. They use a pre-trained AlexNet as a feature extractor. The network includes five convolutional layers and three fully-connected layers, and subsequently feed the extracted feature vectors into an SVM to predict the classification results. Table 11 summarizes the applications of deep learning for segmentation in other types of wounds.

VII. DISCUSSIONS

A. OVERVIEW

From the references in this review, we can find that deep learning has been widely used in various fields of wound image analysis, including the wide range of target wound types and the diversity of processing methods, whereas some of the research results can be reliably applied in clinical practice. We also review the general model frameworks used in the paper (Fig. 2), and find that ResNet, YOLO, and U-Net are the most popular frameworks for classification, detection, and segmentation tasks, respectively. At the same time, during the review, we find that some studies used few general indicators, and some models were only evaluated in unpublished datasets, making it difficult to accurately evaluate their model performance for comparisons. Fortunately, as research on deep learning in wound images continues, and more and more datasets become public, the evaluation of the models will become more thorough and comprehensive. One of the many successful aspects of deep learning is its excellent accuracy. Among them, effective preprocessing methods have contributed significantly, including generating reliable simulated images to expand the training set, reducing optical noise in images, enhancing color information in images, etc. In the study of wound tissues [32], the segmentation accuracy of the combined model trained on images enhanced with color data has increased from 53.83% to 77.65% against necrosis. Model fine-tuning is another method that can effectively improve performance. That is to initialize the existing model with the original model parameters for a small amount of training. Training with pre-trained model parameters can reduce training time and improve the robustness of the model [106]. Then fine-tuning for the task characteristics can lead to better image features. In the experimental results of [81], the classification performance of the fine-tuned model has improved from 76.8% to 84.6% on the wound depth dataset. In addition, choosing a suitable network structure [74], properly adjusting the learning rate [84], and post-processing the model results [28] have been shown to be effective in helping to improve the accuracy rate.

B. CHALLENGES

However, there are still some shortcomings and challenges in deep learning for wound image processing. First, many wound types still lack large-scale publicly available datasets. Although data augmentation and the combination of supervised and unsupervised learning can compensate for this to some extent, the final training results are still some distance away from the upper performance limit of the model itself. Therefore, it is necessary to establish a standardized open dataset. Second, manual annotation of wound images is unstable, especially when dealing with complex wound surfaces, which can lead to inaccuracies in calculating model metrics. Independent annotation of images by multiple physicians may be one way to avoid this drawback, but at a much higher labor cost. In addition, in some of the researches, we get a sense that smart devices are starting to gain attention in wound image analysis. However, the large computational cost required to train the model hinders the widespread use of deep learning in practical application scenarios, and only a few studies have ported the results to mobile devices. With the development of compact networks such as MobileNet and ShufflieNet [107], this challenge may be addressed in the near future.

C. FUTURE RESEARCH DIRECTION

One of the future directions of deep learning in wound research is to increase the number of prediction categories, including wound severity class, tissue class, healing time, etc., so as to extract richer wound information and effectively improve the reliability of clinical decision making. Building a more comprehensive public dataset is another direction of work. The considerations of more types of wound data can enable deep learning methods to be generalized for different types of wounds. The image capture device, the body part where the wound is located, the healing stage of the wound, and the skin color differences among different ethnic groups are also some of the issues to be considered in the data collection, as more data are needed to support the testing of the generalization ability of the model. Considering that the ultimate goal of wound image research is to be reliably applied to real-world needs such as diagnosis, prognosis, and care, how to improve the performance and efficiency of the model is the direction of deep learning in wound image research.

VIII. CONCLUSION

We studied various current applications and recent advances in deep learning applied to wound images, introduced models in different tasks, reviewed publicly available datasets and summarize various data preprocessing approaches. Deep learning is still one of the active research areas in the field of wound image analysis. The reliable performance of deep learning in image classification, detection, and segmentation can effectively improve the diagnostic efficiency of healthcare professionals. For areas with underdeveloped medical resources, it can compensate to a certain extent for the impact caused by the shortage of doctors. With the extensive cooperation of medical institutions, the rapid development of computer hardware and image acquisition equipment, and the continuous optimization of deep learning algorithms, deep learning has a very promising prospect in the field of wound image analysis.

REFERENCES

[1] D. Voegeli, J. Posnett, P. Franks, K. Harding, M. Edmonds, C. Moffatt, and M. Clark, "Skin breakdown: The silent epidemic," Smith Nephew Found., Hull, U.K., Tech. Rep., 2007.
[2] K. Bechert and S. E. Abraham, "Pain management and wound care," *J. Amer. College Certified Wound Specialists*, vol. 1, no. 2, pp. 65–71, 2009.
[3] D. Upton, "Pain, wound care and psychology: The missing link," *Wounds U.K.*, vol. 7, no. 2, pp. 22–119, 2011.
[4] J. F. Guest, N. Ayoub, T. McIlwraith, I. Uchegbu, A. Gerrish, D. Weidlich, K. Vowden, and P. Vowden, "Health economic burden that different wound types impose on the U.K.s national health service," *Int. Wound J.*, vol. 14, no. 2, pp. 322–330, 2017.
[5] R. C. Britt, L. J. Weireter, and L. D. Britt, "Initial implementation of an acute care surgery model: Implications for timeliness of care," *J. Amer. College Surgeons*, vol. 209, no. 4, pp. 421–424, 2009.
[6] R. F. Cubas, N. R. Gómez, S. Rodriguez, M. Wanis, A. Sivanandam, and C. A. Garberoglio, "Outcomes in the management of appendicitis and cholecystitis in the setting of a new acute care surgery service model: Impact on timing and cost," *J. Amer. College Surgeons*, vol. 215, no. 5, pp. 715–721, 2012.
[7] T. E. Serena, J. R. Hanft, and R. Snyder, "The lack of reliability of clinical examination in the diagnosis of wound infection: Preliminary communication," *Int. J. Lower Extremity Wounds*, vol. 7, no. 1, pp. 32–35, Mar. 2008.
[8] S. Monstrey, H. Hoeksema, J. Verbelen, A. Pirayesh, and P. Blondeel, "Assessment of burn depth and burn wound healing potential," *Burns*, vol. 34, no. 6, pp. 761–769, Sep. 2008.
[9] J. Faria, J. Almeida, M. J. M. Vasconcelos, and L. Rosado, "Automated mobile image acquisition of skin wounds using real-time deep neural networks," in *Proc. Annu. Conf. Med. Image Understand. Anal.* Cham, Switzerland: Springer, 2019, pp. 61–73.
[10] H.-H. Tsai, Y.-P. Pong, C.-C. Liang, P.-Y. Lin, and C.-H. Hsieh, "Teleconsultation by using the mobile camera phone for remote management of the extremity wound: A pilot study," *Ann. Plastic Surg.*, vol. 53, no. 6, pp. 584–587, Dec. 2004.
[11] N. Sikka, K. N. Carlin, J. Pines, M. Pirri, R. Strauss, and F. Rahimi, "The use of mobile phones for acute wound care: Attitudes and opinions of emergency department patients," *J. Health Commun.*, vol. 17, no. sup1, pp. 37–43, 2012.
[12] M. Bloemen, P. Van Zuijlen, and E. Middelkoop, "Reliability of subjective wound assessment," *Burns*, vol. 37, no. 4, pp. 566–571, 2011.
[13] F. Veredas, H. Mesa, and L. Morente, "Binary tissue classification on wound images with neural networks and Bayesian classifiers," *IEEE Trans. Med. Imag.*, vol. 29, no. 2, pp. 410–427, Feb. 2009.
[14] R. Mukherjee, D. D. Manohar, D. K. Das, A. Achar, A. Mitra, and C. Chakraborty, "Automated tissue classification framework for reproducible chronic wound assessment," *BioMed Res. Int.*, vol. 2014, pp. 1–9, Jul. 2014.
[15] H. Wannous, Y. Lucas, and S. Treuillet, "Enhanced assessment of the wound-healing process by accurate multiview tissue classification," *IEEE Trans. Med. Imag.*, vol. 30, no. 2, pp. 315–326, Feb. 2011.
[16] F. J. Veredas, R. M. Luque-Baena, F. J. Martín-Santos, J. C. Morilla-Herrera, and L. Morente, "Wound image evaluation with machine learning," *Neurocomputing*, vol. 164, pp. 112–122, Sep. 2015.
[17] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.
[18] D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," *Annu. Rev. Biomed. Eng.*, vol. 19, pp. 221–248, Jul. 2017.
[19] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017.

[20] C. Wang, X. Yan, M. Smith, K. Kochhar, M. Rubin, S. M. Warren, J. Wrobel, and H. Lee, "A unified framework for automatic wound segmentation and analysis with deep convolutional neural networks," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 2415–2418.
[21] D. S. Kermany *et al.*, "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018.
[22] S. Zahia, M. B. G. Zapirain, X. Sevillano, A. González, P. J. Kim, and A. Elmaghraby, "Pressure injury image analysis with machine learning techniques: A systematic review on previous and possible future methods," *Artif. Intell. Med.*, vol. 102, Jan. 2020, Art. no. 101742.
[23] D. M. Anisuzzaman, C. Wang, B. Rostami, S. Gopalakrishnan, J. Niezgoda, and Z. Yu, "Image-based artificial intelligence in wound assessment: A systematic review," *Adv. Wound Care*, Dec. 2021.
[24] S. Thomas. *Medetec Wound Database*. Accessed: Apr. 19, 2022. [Online]. Available: http://www.medetec.co.uk/files/medetec-image-databases.html
[25] J. Chae, K. Y. Hong, and J. Kim, "A pressure ulcer care system for remote medical assistance: Residual U-Net with an attention model based for wound area segmentation," 2021, *arXiv:2101.09433*.
[26] B. García-Zapirain, M. Elmogy, A. El-Baz, and A. S. Elmaghraby, "Classification of pressure ulcer tissues with 3D convolutional neural network," *Med. Biol. Eng. Comput.*, vol. 56, no. 12, pp. 2245–2258, 2018.
[27] F. Li, C. Wang, Y. Peng, Y. Yuan, and S. Jin, "Wound segmentation network based on location information enhancement," *IEEE Access*, vol. 7, pp. 87223–87232, 2019.
[28] C. Wang, D. M. Anisuzzaman, V. Williamson, M. K. Dhar, B. Rostami, J. Niezgoda, S. Gopalakrishnan, and Z. Yu, "Fully automatic wound segmentation with deep convolutional neural networks," *Sci. Rep.*, vol. 10, no. 1, pp. 1–9, Dec. 2020.
[29] A. Mahbod, G. Schaefer, R. Ecker, and I. Ellinger, "Automatic foot ulcer segmentation using an ensemble of convolutional neural networks," 2021, *arXiv:2109.01408*.
[30] S. R. Oota, V. Rowtula, S. Mohammed, J. Galitz, M. Liu, and M. Gupta, "HealTech—A system for predicting patient hospitalization risk and wound progression in old patients," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, Jan. 2021, pp. 2463–2472.
[31] C. Pathompatai, R. Kanawong, and P. Taeprasartsit, "Region-focus training: Boosting accuracy for deep-learning image segmentation," in *Proc. 16th Int. Joint Conf. Comput. Sci. Softw. Eng. (JCSSE)*, Jul. 2019, pp. 319–323.
[32] N. Pholberdee, C. Pathompatai, and P. Taeprasartsit, "Study of chronic wound image segmentation: Impact of tissue type and color data augmentation," in *Proc. 15th Int. Joint Conf. Comput. Sci. Softw. Eng. (JCSSE)*, Jul. 2018, pp. 1–6.
[33] S. Zahia, B. Garcia-Zapirain, and A. Elmaghraby, "Integrating 3D model representation for an accurate non-invasive assessment of pressure injuries with deep learning," *Sensors*, vol. 20, no. 10, p. 2933, May 2020.
[34] B. Rostami, D. M. Anisuzzaman, C. Wang, S. Gopalakrishnan, J. Niezgoda, and Z. Yu, "Multiclass wound image classification using an ensemble deep CNN-based classifier," *Comput. Biol. Med.*, vol. 134, Jul. 2021, Art. no. 104536. [Online]. Available: https://github.com/uwm-bigdata/wound_classification
[35] J. Zhang, E. Zhu, X. Guo, H. Chen, and J. Yin, "Chronic wounds image generator based on deep convolutional generative adversarial networks," in *Proc. Nat. Conf. Theor. Comput. Sci.* New York, NY, USA: Springer, 2018, pp. 150–158.
[36] Burns BIP_US Database. *Biomedical Image Processing (BIP) Group from the Signal Theory and Communications Department (University of Seville, SPAIN) and Virgen del Rocío Hospital (Seville, SPAIN)*. Accessed: Apr. 19, 2022. [Online]. Available: http://personal.us.es/rboloix/Burns_BIP_US_database.zip
[37] R. M. Bhansali and R. Kumar, "BurnNet: An efficient deep learning framework for accurate dermal burn classification," *medRxiv*, Feb. 2021. [Online]. Available: https://www.medrxiv.org/content/10.1101/2021.01.30.21250727v1, doi: 10.1101/2021.01.30.21250727.
[38] B. Rostami, J. Niezgoda, S. Gopalakrishnan, and Z. Yu, "Multiclass burn wound image classification using deep convolutional neural networks," 2021, *arXiv:2103.01361*.
[39] *Sårwebben*. Accessed: Apr. 19, 2022. [Online]. Available: https://www.vgregion.se/s/skaraborgs-sjukhus/vardgivare/sarwebben
[40] C. Aguirre Nilsson and M. Velic, "Classification of ulcer images using convolutional neural networks," M.S. thesis, Dept. Elect. Eng., Chalmers Univ. Technol., Gothenburg, Sweden, 2018.
[41] M. Kręcichwost, J. Czajkowska, A. Wijata, J. Juszczyk, B. Pyciński, M. Biesok, M. Rudzki, J. Majewski, J. Kostecki, and E. Pietka, "Chronic wounds multimodal image database," *Comput. Med. Imag. Graph.*, vol. 88, Mar. 2021, Art. no. 101844. [Online]. Available: https://chronicwounddatabase.eu
[42] B. Monroy, J. Bacca, K. Sanchez, H. Arguello, and S. Castillo, "Two-step deep learning framework for chronic wounds detection and segmentation: A case study in Colombia," in *Proc. 23rd Symp. Image, Signal Process. Artif. Vis. (STSIVA)*, Sep. 2021, pp. 1–6.
[43] M. H. Hesamian, W. Jia, X. He, and P. Kennedy, "Deep learning techniques for medical image segmentation: Achievements and challenges," *J. Digit. Imag.*, vol. 32, no. 4, pp. 582–596, 2019.
[44] M. I. Razzak, S. Naz, and A. Zaib, "Deep learning for medical image processing: Overview, challenges and the future," *Classification BioApps*, vol. 26, pp. 323–350, Nov. 2018.
[45] S. P. Singh, L. Wang, S. Gupta, H. Goli, P. Padmanabhan, and B. Gulyás, "3D deep learning on medical images: A review," *Sensors*, vol. 20, no. 18, p. 5097, 2020.
[46] J. Wang, H. Zhu, S.-H. Wang, and Y.-D. Zhang, "A review of deep learning on medical image analysis," *Mobile Netw. Appl.*, vol. 26, no. 1, pp. 351–380, 2021.
[47] F. Altaf, S. M. Islam, N. Akhtar, and N. K. Janjua, "Going deep in medical image analysis: Concepts, methods, challenges, and future directions," *IEEE Access*, vol. 7, pp. 99540–99572, 2019.
[48] N. Ohura, R. Mitsuno, M. Sakisaka, Y. Terabe, Y. Morishige, A. Uchiyama, T. Okoshi, I. Shinji, and A. Takushima, "Convolutional neural networks for wound detection: The role of artificial intelligence in wound care," *J. Wound Care*, vol. 28, no. 10, pp. S13–S24, Oct. 2019.
[49] V. Rajathi, R. R. Bhavani, and G. W. Jiji, "Varicose ulcer(C6) wound image tissue classification using multidimensional convolutional neural networks," *Imag. Sci. J.*, vol. 67, no. 7, pp. 374–384, Oct. 2019.
[50] A. Wagh, S. Jain, A. Mukherjee, E. Agu, P. C. Pedersen, D. Strong, B. Tulu, C. Lindsay, and Z. Liu, "Semantic segmentation of smartphone wound images: Comparative analysis of AHRF and CNN-based approaches," *IEEE Access*, vol. 8, pp. 181590–181604, 2020.
[51] 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 *Proc. IEEE 28th Int. Workshop Mach. Learn. Signal Process. (MLSP)*, Sep. 2018, pp. 1–6.
[52] F. Dai, D. Zhang, K. Su, and N. Xin, "Burn images segmentation based on burn-GAN," *J. Burn Care Res.*, vol. 42, no. 4, pp. 755–762, Aug. 2021.
[53] G. Ranganathan, "A study to find facts behind preprocessing on deep learning algorithms," *J. Innov. Image Process.*, vol. 3, no. 1, pp. 66–74, Apr. 2021.
[54] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and D. L. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Comput.*, vol. 1, no. 4, pp. 541–551, 1989.
[55] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.
[56] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[57] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, Apr. 2017.
[58] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.
[59] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.
[60] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4700–4708.
[61] E. P. Ong, C. T. K. Yin, and B.-H. Lee, "Efficient deep learning-based wound-bed segmentation for mobile applications," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 1654–1657.
[62] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 580–587.
[63] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 1440–1448.

[64] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 28, 2015, pp. 1–9.
[65] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 779–788.
[66] J. Du, "Understanding of object detection based on CNN family and YOLO," *J. Phys., Conf.*, vol. 1004, Apr. 2018, Art. no. 012029.
[67] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot multibox detector," in *Proc. Eur. Conf. Comput. Vis.* New York, NY, USA: Springer, 2016, pp. 21–37.
[68] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, "End-to-end object detection with transformers," in *Proc. Eur. Conf. Comput. Vis.* New York, NY, USA: Springer, 2020, pp. 213–229.
[69] A. Oliveira, A. Britto de Carvalho, and D. Dantas, "Faster R-CNN approach for diabetic foot ulcer detection," in *Proc. 16th Int. Joint Conf. Comput. Vis., Imag. Comput. Graph. Theory Appl.*, 2021, pp. 677–684.
[70] J. Amin, M. Sharif, M. A. Anjum, H. U. Khan, M. S. A. Malik, and S. Kadry, "An integrated design for classification and localization of diabetic foot ulcer based on CNN and YOLOv2-DFU models," *IEEE Access*, vol. 8, pp. 228586–228597, 2020.
[71] F. A. Khan, A. U. R. Butt, M. Asif, W. Ahmad, M. Nawaz, M. Jamjoom, and E. Alabdulkreem, "Computer-aided diagnosis for burnt skin images using deep convolutional neural network," *Multimedia Tools Appl.*, vol. 79, nos. 45–46, pp. 34545–34568, Dec. 2020.
[72] Y. Wang, Z. Ke, Z. He, X. Chen, Y. Zhang, P. Xie, T. Li, J. Zhou, F. Li, C. Yang, P. Zhang, C. Huang, and L. Kai, "Real-time burn depth assessment using artificial networks: A large-scale, multicentre study," *Burns*, vol. 46, no. 8, pp. 1829–1838, Dec. 2020.
[73] J. Chauhan and P. Goyal, "BPBSAM: Body part-specific burn severity assessment model," *Burns*, vol. 46, no. 6, pp. 1407–1423, Sep. 2020.
[74] M. D. Cirillo, R. Mirdell, F. Sjöberg, and T. D. Pham, "Time-independent prediction of burn depth using deep convolutional neural networks," *J. Burn Care Res.*, vol. 40, no. 6, pp. 857–863, Oct. 2019.
[75] A. Abubakar, H. Ugail, and A. M. Bukar, "Assessment of human skin burns: A deep transfer learning approach," *J. Med. Biol. Eng.*, vol. 40, no. 3, pp. 321–333, Jun. 2020.
[76] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 3431–3440.
[77] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput.-Assist. Intervent*. New York, NY, USA: Springer, 2015, pp. 234–241.
[78] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 2961–2969.
[79] C. Jiao, K. Su, W. Xie, and Z. Ye, "Burn image segmentation based on mask regions with convolutional neural network deep learning framework: More accurate and more convenient," *Burns Trauma*, vol. 7, pp. 1–14, Dec. 2019.
[80] M. Hossin and M. N. Sulaiman, "A review on evaluation metrics for data classification evaluations," *Int. J. Data Mining Knowl. Manage. Process*, vol. 5, no. 2, pp. 1–11, Mar. 2015.
[81] X. Zhao, Z. Liu, E. Agu, A. Wagh, S. Jain, C. Lindsay, B. Tulu, D. Strong, and J. Kan, "Fine-grained diabetic wound depth and granulation tissue amount assessment using bilinear convolutional neural network," *IEEE Access*, vol. 7, pp. 179151–179162, 2019.
[82] C. Gamage, I. Wijesinghe, and I. Perera, "Automatic scoring of diabetic foot ulcers through deep CNN based feature extraction with low rank matrix factorization," in *Proc. IEEE 19th Int. Conf. Bioinf. Bioengineering (BIBE)*, Oct. 2019, pp. 352–356.
[83] L. Alzubaidi, M. A. Fadhel, S. R. Oleiwi, O. Al-Shamma, and J. Zhang, "DFU_QUTNet: Diabetic foot ulcer classification using novel deep convolutional neural network," *Multimedia Tools Appl.*, vol. 79, nos. 21–22, pp. 15655–15677, Jun. 2020.
[84] S. K. Das, P. Roy, and A. K. Mishra, "DFU_SPNet: A stacked parallel convolution layers based CNN to improve diabetic foot ulcer classification," *ICT Exp.*, vol. 8, no. 2, pp. 271–275, Jun. 2021.
[85] 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 Trans. Emerg. Topics Comput. Intell.*, vol. 4, no. 5, pp. 728–739, Sep. 2018.
[86] M. Goyal, N. D. Reeves, S. Rajbhandari, N. Ahmad, C. Wang, and M. H. Yap, "Recognition of ischaemia and infection in diabetic foot ulcers: Dataset and techniques," *Comput. Biol. Med.*, vol. 117, Feb. 2020, Art. no. 103616.
[87] G. Blanco, A. J. M. Traina, C. Traina, P. M. Azevedo-Marques, A. E. S. Jorge, D. de Oliveira, and M. V. N. Bedo, "A superpixel-driven deep learning approach for the analysis of dermatological wounds," *Comput. Methods Programs Biomed.*, vol. 183, Jan. 2020, Art. no. 105079.
[88] R. Brungel and C. M. Friedrich, "DETR and YOLOv5: Exploring performance and self-training for diabetic foot ulcer detection," in *Proc. IEEE 34th Int. Symp. Computer-Based Med. Syst. (CBMS)*, Jun. 2021, pp. 148–153.
[89] A. Han, Y. Zhang, A. Li, C. Li, F. Zhao, Q. Dong, Q. Liu, Y. Liu, X. Shen, S. Yan, and S. Zhou, "Efficient refinements on YOLOv3 for real-time detection and assessment of diabetic foot wagner grades," 2020, *arXiv:2006.02322*.
[90] B. Cassidy, N. D. Reeves, J. M. Pappachan, N. Ahmad, S. Haycocks, D. Gillespie, and M. H. Yap, "A cloud-based deep learning framework for remote detection of diabetic foot ulcers," *IEEE Pervasive Comput.*, vol. 21, no. 2, pp. 78–86, Apr. 2022.
[91] M. H. Yap *et al.*, "Deep learning in diabetic foot ulcers detection: A comprehensive evaluation," *Comput. Biol. Med.*, vol. 135, Aug. 2021, Art. no. 104596.
[92] J. Chauhan and P. Goyal, "Convolution neural network for effective burn region segmentation of color images," *Burns*, vol. 47, no. 4, pp. 854–862, Jun. 2021.
[93] O. Despo, S. Yeung, J. Jopling, B. Pridgen, C. Sheckter, S. Silberstein, L. Fei-Fei, and A. Milstein, "BURNED: Towards efficient and accurate burn prognosis using deep learning," Stanford Univ., Palo Alto, CA, USA, Tech. Rep., 2017.
[94] H. Liu, K. Yue, S. Cheng, W. Li, and Z. Fu, "A framework for automatic burn image segmentation and burn depth diagnosis using deep learning," *Comput. Math. Methods Med.*, vol. 2021, pp. 1–12, Apr. 2021.
[95] M.-S. Badea, C. Vertan, C. Florea, L. Florea, and S. Badoiu, "Automatic burn area identification in color images," in *Proc. Int. Conf. Commun. (COMM)*, Jun. 2016, pp. 65–68.
[96] C. Pabitha and B. Vanathi, "Densemask RCNN: A hybrid model for skin burn image classification and severity grading," *Neural Process. Lett.*, vol. 53, no. 1, pp. 319–337, Feb. 2021.
[97] M. D. Cirillo, R. Mirdell, F. Sjöberg, and T. D. Pham, "Improving burn depth assessment for pediatric scalds by AI based on semantic segmentation of polarized light photography images," *Burns*, vol. 47, no. 7, pp. 1586–1593, Nov. 2021.
[98] C. Cui, K. Thurnhofer-Hemsi, R. Soroushmehr, A. Mishra, J. Gryak, E. Domínguez, K. Najarian, and E. López-Rubio, "Diabetic wound segmentation using convolutional neural networks," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 1002–1005.
[99] P. Muñoz, R. Rodríguez, and N. Montalvo, "Automatic segmentation of diabetic foot ulcer from mask region-based convolutional neural networks," *J. Biomed. Res. Clin. Invest.*, vol. 2, no. 1, pp. 1–9, Aug. 2020.
[100] N. Rania, H. Douzi, L. Yves, and T. Sylvie, "Semantic segmentation of diabetic foot ulcer images: Dealing with small dataset in DL approaches," in *Proc. Int. Conf. Image Signal Process.* New York, NY, USA: Springer, 2020, pp. 162–169.
[101] R. Niri, H. Douzi, Y. Lucas, and S. Treuillet, "A superpixel-wise fully convolutional neural network approach for diabetic foot ulcer tissue classification," in *Proc. Int. Conf. Pattern Recognit.* New York, NY, USA: Springer, 2021, pp. 308–320.
[102] C. W. Chang, M. Christian, D. H. Chang, F. Lai, T. J. Liu, Y. S. Chen, and W. J. Chen, "Deep learning approach based on superpixel segmentation assisted labeling for automatic pressure ulcer diagnosis," *PLoS ONE*, vol. 17, no. 2, Feb. 2022, Art. no. e0264139.
[103] D. Y. T. Chino, L. C. Scabora, M. T. Cazzolato, A. E. S. Jorge, C. Traina-Jr., and A. J. M. Traina, "Segmenting skin ulcers and measuring the wound area using deep convolutional networks," *Comput. Methods Programs Biomed.*, vol. 191, Jul. 2020, Art. no. 105376.
[104] H. Gamage, W. Wijesinghe, and I. Perera, "Instance-based segmentation for boundary detection of neuropathic ulcers through mask-RCNN," in *Proc. Int. Conf. Artif. Neural Netw.* New York, NY, USA: Springer, 2019, pp. 511–522.
[105] H. Nejati, H. A. Ghazijahani, M. Abdollahzadeh, T. Malekzadeh, N.-M. Cheung, K.-H. Lee, and L.-L. Low, "Fine-grained wound tissue analysis using deep neural network," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2018, pp. 1010–1014.
[106] D. Hendrycks, K. Lee, and M. Mazeika, "Using pre-training can improve model robustness and uncertainty," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 2712–2721.
[107] X. Zhang, X. Zhou, M. Lin, and J. Sun, "ShuffleNet: An extremely efficient convolutional neural network for mobile devices," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 6848–6856.