BMC Medical Imaging

RESEARCH

Open Access

Eff-ReLU-Net: a deep learning framework for multiclass wound classification

Sifat Ullah[1], Ali Javed[1], Muteb Aljasem[2] and Abdul Khader Jilani Saudagar[3*]

Abstract

Chronic wounds have emerged as a significant medical challenge due to their adverse effects, including infections leading to amputations. Over the past few years, the prevalence of chronic wounds has grown, thus posing significant health hazards. It is now becoming necessary to automate the wound assessment mechanism to limit the dependence of healthcare practitioners on manual methods. Therefore, a need exists for developing an effective wound classifier that enables practitioners to classify wounds quickly and reliably. This work proposed Eff-ReLU-Net, an improved EfficientNet-B0-based deep learning model for accurately identifying multiple categories of wounds. More precisely, we introduced the ReLU activation function over the Swish in our Eff-ReLU-Net because of its simplicity, reliability, and efficiency. Additionally, we introduced three fully connected dense layers at the end to reliably capture more distinct features, leading to improved multi-class wound classification. We also employed augmentation approaches such as fixed-angle rotations at 90°, 180°, and 270°, rotational invariance, random rotation, and translation to improve data diversity and samples for better model generalization and combating overfitting. The proposed model's effectiveness is assessed utilizing the publicly available AZH and Medetec wound datasets. We also conducted the cross-corpora evaluation to show the generalizability of our method. The proposed model achieved an accuracy, precision, recall, and F1-score of 92.33%, 97.66%, 95.33%, and 96.48% on Medetec, respectively. However, for the AZH dataset, the attained accuracy, precision, recall, and F1-score are 90%, 89.45%, 92,19%, and 90.84%, respectively. These results validate the effectiveness of our proposed Eff-ReLU-Net method for classifying chronic wounds.

Keywords Chronic wound classification, Eff-ReLU-Net, EfficientNet, Medetec, Rectified learning unit

Introduction

Wounds are injuries that affect the skin and underlying tissues and are mainly caused by disease, surgical procedures, or trauma. Wounds are generally classified as chronic or acute based on the extent of the physical damage and the timeframe from the primary injury. Compared to acute wounds, which often improve in four weeks, chronic wounds resist recovery within six weeks. Chronic wounds are critical and require appropriate medication, as such wounds are severe, painful, tender, and can cause significant problems for patients. According to a research study [1], yearly, over 6 million Americans suffer from chronic wounds. Diabetic foot ulcers

*Correspondence:
Abdul Khader Jilani Saudagar
aksaudagar@imamu.edu.sa
[1]Department of Software Engineering, University of Engineering and Technology-Taxila, Taxila 47050, Pakistan
[2]Department of Robotics, Electronics, and Computer Engineering, Bowling Green State University, Bowling Green, OH, USA
[3]Information Systems Department, College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11432, Saudi Arabia

(DFU), venous leg ulcers (VLU), and pressure ulcers (PU) are the most common types of chronic wounds. These chronic wounds impose a substantial burden on the healthcare system as they need continued monitoring and specialized treatment. Thus, medical costs and resource utilization significantly increase due to prolonged treatments and frequent hospital visits [2].

Conventionally, practitioners follow manual wound assessment methods for treating and caring for chronic wounds. However, manual wound assessment can be inconsistent and subjective, leading to variability in diagnoses and treatment plans. This highlights the need for automated wound classification systems to provide reliable and objective evaluations to support clinicians in delivering effective patient care [3]. Scholars have presented various conventional machine learning (ML) and deep learning (DL) based approaches for automatic wound classification. ML approaches learn hand-crafted features, capturing attributes like texture, color, and shape from the wound images, and require domain-specific expertise for effective feature selection [4, 5]. Such methods are ineffective for learning complex patterns in wound images and have poor generalizability across diverse wound types [6, 7]. On the other hand, DL-based approaches have improved wound classification via learning hierarchical and complex features from the input wound images [2, 8–10]. Deep learning networks have been proven effective for improving wound diagnosis, including wound classification. However, DL-based approaches usually need large datasets to achieve enhanced performance and more processing power [11].

Most of the existing methods have tackled wound identification as a binary classification problem, showing the presence or absence of wounds. However, automated methods must also be capable of identifying the type of wounds to provide more useful information to medical surgeons. In recent times, researchers have also developed multiclass wound classification methods to indicate the type of wound. Despite progress in deep learning methods, existing wound classification systems still struggle with accurate identification across multiple wound types, especially due to variations in the color, shape, and size of wounds. Moreover, wounds that are about to heal are particularly challenging to detect. These challenges are further compounded by limited datasets and computational inefficiencies in model design, which hinder the development of robust and generalizable classification systems.

To overcome current limitations, this study focuses on developing an efficient and accurate multiclass chronic wound classification model that generalizes well across different datasets. For this study, we formulated two research questions: (1) Can the use of ReLU activation in place of Swish improve the efficiency and training stability of the wound classification model? and (2) Can the proposed model generalize well across datasets with different wound types? For this, we introduced the computationally efficient ReLU activation function over the swish activation to enhance the efficiency of the proposed wound classification. ReLU outputs non-negative values by avoiding complex exponential computations, converging faster, and preventing the vanishing gradient problem. These attributes enable ReLU to become very efficient in deep neural networks, thus refining training proficiency and efficiency. Further, we also suggested increasing the fully connected dense layers for more effective feature computation. Additionally, we applied data augmentation approaches to enrich the dataset, reducing overfitting and enhancing generalization without requiring significant computational resources. Along with that, cross-corpora evaluation is conducted to analyze the generalization aptitude of the proposed model. The main contributions of the proposed research are:

- We present an Eff-ReLU-Net method to reliably classify multiple categories of chronic wounds, irrespective of diverse wound types and sizes.
- We introduce the ReLU activation function in our Eff-ReLU-Net to improve the efficiency and reliability of the baseline EfficientNet-B0 model.
- We introduce three fully connected dense layers at the end to better capture the distinct traits of different types of wound images.
- Extensive experimentation was conducted, including cross-dataset evaluation using Medetec and AZH datasets, to demonstrate the proposed model's superiority and generalizability over existing approaches.

The remaining paper is organized as follows. The related work is reviewed in Sect. 2. Section 3 presents the suggested approach, experimental setup, utilized datasets, data augmentation, and evaluation metrics. Experimental results are discussed in Sect. 4. Finally, Sects. 5 and 6 demonstrate the analysis and conclusion of our work.

Related work

In this section, we discussed recent work on wound image classification using traditional ML and DL-based methods.

Traditional handcrafted methods of wound classification mainly depend on the health practitioners' knowledge. They consider features like wound location, size, depth, appearance of wound edges, and surrounding skin condition, frequently using basic tools like rulers and probes to access key features with classifications. Initially, conventional approaches were used to recognize wounds for binary classification, like SVM and decision trees, to

distinguish wound and non-wound regions in medical images, including typical skin or background elements. Examples of binary wound type classification scenarios involve classifying background from DFU, normal skin from PU, and DFU from PU. Yadav et al. [12] proposed a traditional machine learning method for binary classification of burn wound images using a classic color-based feature extraction technique with an SVM. This method [12] categorized the samples into two classes: grafting and non-grafted wounds, while achieving 82.43% accuracy. The dataset comprised 94 images with burn depths, including full-thickness, deep dermal, and superficial dermal. This method was evaluated on a small custom dataset, and its assessment on a large and diverse dataset is yet to be explored. Multi-class image classification using conventional ML methods is adopted by different studies like an approach discussed in [13], where the Scale Invariant Feature Transform (SIFT) and Speeded Robust Features (SURF) methods combined with the Bag of Features (BOF) technique are used to extract the features from given images to train the SVM for classification. The approach attained a classification score of 97.81%, however, it needs testing on a large and diverse dataset, where a significant reduction in performance is expected. Veredas et al. [14] used the three machine learning models: neural networks, support vector machines (SVMs), and random forest decision trees to categorize every segmented area of wound tissue. Models were trained to classify tissue categories built on features extracted from the segmented regions, with the neural network attaining an accuracy of 81.87%, SVMs achieving 88.08%, and random forests getting 87.37%. Alevizos et al. [15] proposed an ML-based approach for recognizing diabetic foot ulcers. For this, the approach suggested an improved local binary pattern (LBP) to compute rotation-invariant feature vectors. The estimated information is later categorized by using the XGBoost approach. The work has attained a sensitivity score of 92%, however, it needs more improvements. ML-based approaches for wound classification perform well, however, they often face challenges such as reliance on manual feature extraction, which can be time-consuming and error-prone, and limited generalization due to insufficient dataset diversity. Additionally, these methods struggle with handling variations in wound size, texture, and imaging conditions, leading to inconsistent performance in real-world scenarios. As a result, traditional ML models may not accurately classify wounds, leading to biased decisions and limited clinical utility under scenarios where wound samples possess significant variations in size, color, and shape.

The advancement in deep learning approaches has shifted the paradigm, and various models like pre-trained CNNs are widely explored for medical image analysis, including wound classification. Goyal et al. [16] classify the infection and ischemia in diabetic wounds using an ensemble CNN network for binary classification. A custom DFU dataset was utilized in his research, and lower accuracy was achieved for the infected class. Rostami et al. [17] classify wound images into binary and multi-classes by using an end-to-end ensemble DCNN-based model utilizing the Medetec dataset. The work reported accuracy values of 91.2% and 82.9% for binary and 3-class classifications. The method performs well for binary classification; however, it needs improvements for multi-class classification. In addition, Husers et al. [18] used YOLOv5 architecture to classify DFU and VU, with the YOLOv5m6 architecture attaining 94.2% accuracy on a custom dataset of 885 images. This study is tested on a small dataset. Husers et al. [19] classify macerations in diabetic foot ulcers by using a Mobilenet-v1-based model utilizing the transfer learning method. This method was tested on a small custom dataset of 416 wound images collected from the wound care center of Christliches Klinikum Melle, Germany. The work reported an accuracy of 69%, which needs significant improvement. Aguirre et al. [20] used a pre-trained VGG-19 model to categorize the ulcer images and attained 75% accuracy and 82% precision for the custom dataset of 300 ulcer images. Ahsan et al. [21] used VGG16, ResNet50, and AlexNet, utilizing the DFU2020 dataset, attaining an accuracy of 99% for ischemia and 84.76% for infection classification on ResNet50. Rostami et al. [22] classify wound images into multiple classes, including surgical, diabetic, and venous ulcers, by using an end-to-end ensemble DCNN-based model utilizing the Medetec dataset. The output classification values of two classifiers depend on patch-wise and image-wise approaches that were input into a Multi-Layer Perceptron to give a larger classifier. The work reported 91.9% and 87.7% accuracy for three classes and 96.4% and 94.28% for binary class classification. Sarp et al. [23] classify chronic wounds into diabetic, lymph vascular, pressure injury, and surgical wounds using an explainable artificial intelligence method adopting the dataset of the eKare repository, which contains 8690 wound images. The VGG16 architecture was used for the classification. The work reported an average F1 score of 0.76. Zaid et al. [24] proposed a Swish-ELU EfficientNet-B4 model that merges two modified EfficientNet-B4 models to get better relevant features, which are then classified through a regularized prediction module. The AZH and Medetec datasets were used for testing. This method attained 87.32% and 88.17% accuracy on the AZH and Extended AZH datasets, whereas 88% and 89.34% accuracy on the Medetec and extended Medetec datasets, respectively.

Qian et al. [25] suggested a deep framework called the SARNET in which a self-attention embedded residual

model is applied to compute the features and categorize the wounds into 6 groups. The approach has reported a maximum classification performance of 80%, which needs further enhancement. Patel et al. [26] proposed a multi-modal combined VGG16, ResNet152, and EfficientNet deep convolutional neural architecture for the classification of wounds into four classes: diabetic, pressure, surgical, and venous ulcers, utilizing corresponding body location. This multi-modal architecture was trained and assessed on Medetec and the AZH datasets. Abubakar et al. [27] suggested an approach to differentiate pressure ulcers and burn wounds. Pre-trained deep learning models like ResNet152, ResNet101, and VGG-face were used for feature extraction. While the SVM was used for classification. The dataset comprised 31 burn images and 29 pressure images, which were increased using flipping transformations, rotation, and cropping. The author executed a 3-class classification and binary class tests. The images were categorized into pressure classes or burn images in the binary classification. The 3-class classification test aimed to categorize the images as pressure, burn, or healthy skin. The ResNet152 model performed best among the employed DL models. Alhababi et al. [28] proposed a framework comprising an attention-dense-unit method for wound segmentation and a feature concatenation-based method for wound classification. This unified method attained 90% accuracy on the Medetec dataset and 81% accuracy on the AZH dataset for wound classification. These hybrid methods [26–28] are computationally complex due to employing multiple DL models for wound classification. DFU_MultiNet, introduced in [29], detected diabetic foot ulcers utilizing multiscale feature fusion. DFU_XAI [30] was an end-to-end framework that fine-tuned five pretrained CNNs (Xception, DenseNet121, ResNet50, InceptionV3, MobileNetV2) and applied three XAI methods (SHAP, LIME, Grad-CAM) to both classify and localize ulcers. These frameworks [29, 30] are limited to the DFU-specific dataset and lack generalizability. Likewise [29], Biswas et al. [31] demonstrated a multi-scale feature-fusion network named XAI-FusionNet that fused high-frequency features from DenseNet201, VGG19, and NASNetMobile via a meta-tuner module, and applied SHAP, LIME, and Grad-CAM for post-hoc explainability. FusionNet [31] achieved 99.05% accuracy and 99.09% AUC. However, this study is limited to binary classification, misclassifying other ulcer types, indicating a need for future extensions to multi-class detection to enhance its clinical applicability.

Table 1 Overview of wound classification methods

Ref.	Methods	Accuracy (%)	Dataset	Limitations
[12]	Color descriptor + SVM	82.43	Burns BIP_US Database	Tested on a small custom dataset
[16]	Ensemble CNN network	73–90	DFU dataset	Specific for DFUs
[19]	Mobilenet-v1	69	Custom dataset	Limited data samples are used Lower accuracy
[27]	VGG-face, ResNet101, and ResNet152	99	Custom dataset	Designed for limited datasets
[22]	Ensemble DCNN	82.9	MEDETEC dataset	Not robust for different datasets
[23]	VGG16 architecture	76	eKare, Inc dataset	Low classification accuracy
[17]	Ensemble DCNN	91.20	Medetec dataset	Needs improvements for multi-class classification
[24]	Swish-ELU EfficientNet-B4	89.34	MEDETEC, AZH, and Extended MEDETEC	Computationally inefficient
[20]	VGG-19	75	Custom dataset	Low classification accuracy
[25]	SARNET	80	kaggle dataset	Low classification accuracy
[21]	AlexNet, VGG16, and ResNet50	84.76–99.49	DFU2020	Computational Complex
[18]	YOLOv5	94	Custom dataset	Tested on a small custom dataset
[28]	Att-d-UNet	90	AZH and Medetec	Computationally inefficient
[26]	VGG16, ResNet152, and EfficientNet	78.10	AZH and Medetec	Computational Complex
[13]	SIFT-SURF-BAG + SVM	99.09	Custom dataset	Tested on a small custom dataset
[14]	Random forest, SVM, and decision trees	81-88.07	Custom dataset	Low accuracy
[15]	LBP with XGBoost	76.5	Custom dataset	Low accuracy

Despite the advancements offered by deep learning approaches in addressing the limitations of traditional methods, there remains room for performance improvement. Challenges such as optimizing computational efficiency, enhancing model generalization to diverse and unseen datasets, and improving robustness to imaging variations highlight the need for further research and innovation in this domain. Table 1 shows the comparison of the existing works.

Proposed methodology

This section presents Eff-ReLU-Net, an end-to-end deep-learning model for classifying wound types, including diabetic foot ulcers, pressure wounds, toe ulcers, and leg ulcers. A comprehensive visual demonstration of the proposed approach is shown in Fig. 1. In the preprocessing step, the input image is resized to 224×224 resolution to satisfy the input size specifications of the

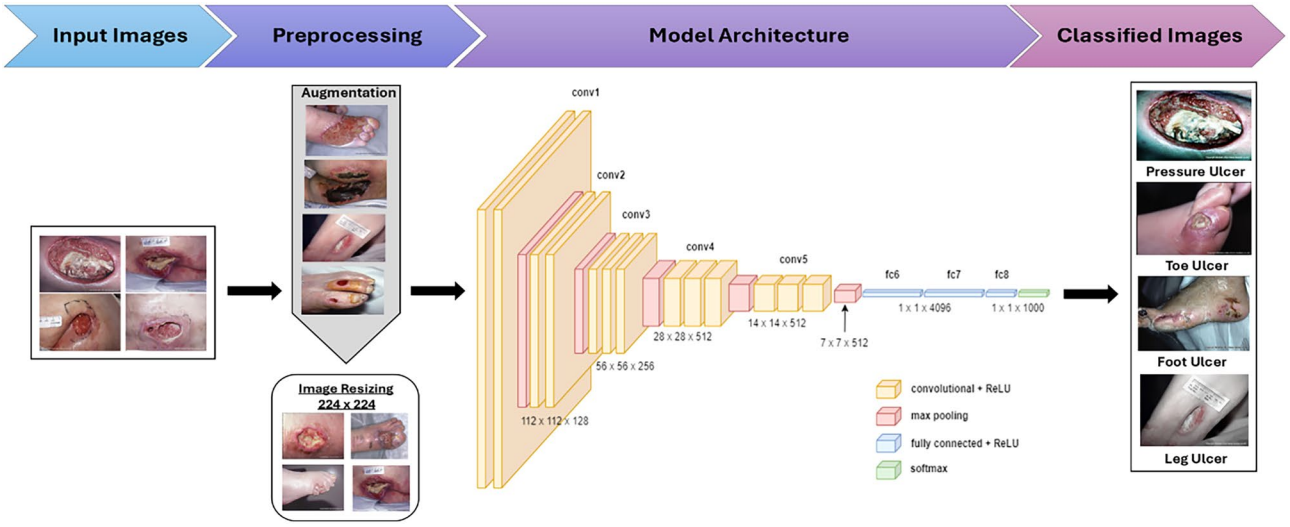


Fig. 1 Proposed Framework

proposed model. Then, the wound images are augmented to increase the number of samples in each wound class, ensuring that the training samples are sufficient to train the proposed deep learning model. The proposed Eff-ReLU-Net model then processes the resized image using multiple convolution blocks, comprising convolution and max pooling layers, along with the ReLU activation function and three fully connected dense layers before the final output layer to perform multi-class wound classification. A detailed description of the proposed model, Eff-ReLU-Net, is provided in the subsequent subsections.

EfficientNet

EfficientNet is a convolutional neural network that employs a compound scaling method to optimize the balance between network depth, width, and resolution, thus permitting it to attain maximum accuracy while sustaining computational efficacy [32]. These attributes make EfficientNet particularly well-suited for tasks that require both precision and resource management, as it consistently outperforms other deep learning models in terms of accuracy-to-efficiency ratio. EfficientNet's ability to capture fine-grained details with reduced computational cost makes it an ideal choice for the wound classification task. EfficientNet-B0 from the EfficientNet family (i.e., EfficientNet B0 to B7) is utilized in our research work, for feature extraction and wound classification, because it has the least number of parameters (5.3 M) and FLOPS (0.39B) on ImageNet, shown in an experiment in [32]. The most accurate model in the EfficientNet family is the EfficientNet-b7, however, it is computationally expensive with 66 M parameters and 37B FLOPS [32]. So, we choose EfficientNet-B0's architecture for customization as it is more cost-effective, robust, and efficient among other models in the EfficientNet family. Moreover, compared to other lightweight models such as MobileNetV2 (3.4 B FLOPS, ~3.5 M parameters) or ShuffleNet (0.14B FLOPS), EfficientNet-B0 matches or exceeds their accuracy because of its optimized inverted-bottleneck blocks with squeeze-and-excitation. Overall, we choose the Efficient-B0 model that not only performs well in classifying wound types but is also useful in real-world clinical environments where resources are limited.

Eff-ReLU-Net architecture

EfficientNet has established itself as a state-of-the-art architecture for image classification tasks that delivers an excellent balance between accuracy and computational efficiency. Its compound scaling method ensures optimized performance across a range of model sizes, which makes it highly versatile. However, despite its strengths, EfficientNet faces certain limitations when applied to complex and domain-specific tasks like wound classification. The default use of the Swish activation function, while beneficial in some contexts, however, can introduce computational overhead, which increases training time, and its dependence on the sigmoid function leads to gradient saturation, slowing convergence. Additionally, Swish struggles with sparse feature representations, which are critical for wound-classification tasks where complex details play a significant role. To address these issues, we have developed the Eff-ReLU-Net model to enhance the accuracy of wound classification. Specifically, we have incorporated the ReLU activation function in EfficientNet. ReLU is simpler and less computationally demanding, which reduces processing time and enhances the suitability of the framework for real-world clinical deployment. Further, this activation function allows the approach to mitigate vanishing gradient problems, enabling faster convergence and improved feature

learning. We include average pooling and dropout layers to avoid overfitting. We also added three fully connected dense layers to further extract and pass the distinctive features to the SoftMax layer to classify different types of wounds. This modification not only reduced training complexity but also enhanced the capability of our approach to extract fine-grained wound features, resulting in higher classification accuracy.

Specifically, the Eff-ReLU-Net model begins with an input layer that processes images resized to 224 × 224 pixels. The architecture incorporates Mobile Inverted Bottleneck Convolutions (MBConvs), which form the backbone of the network. These layers utilize depthwise separable convolutions and inverted residual connections to effectively extract meaningful features while reducing computational overhead. Unlike the original EfficientNet, ReLU is used as the activation function instead of Swish, simplifying computations while retaining performance. These blocks are repeated with varying channel sizes and strides for progressive downsampling and feature enhancement. A final 1 × 1 convolution layer is followed by global average pooling and three fully connected (FC) layers. The last layer accomplishes the classification task using a softmax activation function to produce class probabilities. Table 2 provides the details of the proposed architecture.

Table 2 Details of Eff-ReLU-Net

S. No	Operator	Image Size	Channel	Layers
1	Input Conv 3×3	224×224	32	1
2	MBConv1 3×3 (Normalization + Relu)	112×112	16	1
3	MBConv6 3×3 (Normalization + Relu)	112×112	24	2
4	MBConv6 5×5(Normalization + Relu)	56×56	40	2
5	MBConv6 3×3 (Normalization + Relu)	28×28	80	3
6	MBConv6 5×5 (Normalization + Relu)	14×14	112	3
7	MBConv6 5×5(Normalization + Relu)	14×14	192	4
8	MBConv6 3×3 (Normalization + Relu)	7×7	320	1
9	Conv 1×1, pooling	7×7	1280	1
10	Fully Connected (Dense) Layer 1	1×1	Units: 512	-
11	Fully Connected (Dense) Layer 2	1×1	Units: 256	-
12	Fully Connected (Dense) Layer 3	1×1	Units: 128	-
13	Classification	1×1		-

Convolutional and residual bottleneck layers

The model employs five convolutional layers with a stride of 2 × 2 to reduce spatial resolution. The initial convolutional layer contains 32 kernels and is followed by 15 MBConv layers, which form the core of the architecture. Each MBConv layer includes a depth-wise convolution with a 3 × 3 kernel and a pointwise convolution with a 1 × 1 filter. This configuration is pivotal for capturing fine-grained features while maintaining the model's computational efficiency. After each convolution, a batch normalization (BN) layer is applied to stabilize the learning process, except after the third convolution, where the BN layer directly follows a 1 × 1 convolution.

Activation function and pooling

In the Eff-ReLU-Net model, the ReLU activation function is used with convolutional and fully connected layers to establish non-linearity, which is critical for learning complicated patterns in wound images. We selected ReLU because of its simplicity, computational effectiveness, and capability to prevent vanishing gradient issues, which guarantees speedy training and effective learning even in deeper networks. ReLU's sparse activation also proceeds as implicit regularization, reducing overfitting by forcing the model to focus on the most appropriate features. Unlike other complex functions like Swish, which can be computationally expensive, ReLU allows for faster convergence without degrading performance. Moreover, we integrated a global average pooling layer to ease the model's parameters, improving computational proficiency while preserving the fundamental information needed for precise wound classification.

FC dense and softmax layers

We introduced three FC-dense layers in our Eff-ReLU-Net architecture to extract and pass more distinctive patterns to the SoftMax classification layer to convert the output into probability distributions among the various wound classes. The class with the maximum probability is chosen as the calculated wound class. Statistically, the SoftMax function is stated as:

$$\sigma\ (Zi) = \frac{e^{zi}}{\sum_{j=1}^{n} e^{zj}} \tag{1}$$

Where z_i is the input to the i-th neuron in the output layer, and n is the total number of classes. The SoftMax function exponentiates each z_i, getting all values positive, then normalizes them by dividing with the sum of all exponentiated values. This certifies that $\sigma_{(Zi)}$ denotes the probability of the input image relating to class I, allowing the model to classify the image into predefined wound classes based on the maximum probability. The algorithm

for the classification process of wound images is presented as Algorithm 1.

Algorithm 1: Classification process of proposed Eff-ReLU-Net

Input: *Image Repository*, $X = \{x_1, x_2, x_3, \ldots, x_n\}$
Augmented Image Repository, $W = \{w_1, w_2, w_3, \ldots, w_n\}$
Output: *Predicted class label*, Y_p
1. ***For****each****w****in****W****do***
2. ⟶ *w* ⟵ *Resize (w) // Resized the image to 224 × 224*
3. ⟶ *w* ⟵ *Conv 3 × 3 (w, out_channels = 32)*
4. ⟶ *w* ⟵ *BatchNorm (w)*
5. ⟶ *w* ⟵ *ReLU (w)*
6. ⟶ *w* ⟵ *MBConv (w, exp = 1, out = 16, kernel = 3, stride = 1)*
7. *Repeat 2 times*
8. ⟶ *w* ⟵ *MBConv (w, exp = 6, out = 24, kernel = 3, stride = 1)*
9. *Repeat 2 times*
10. ⟶ *w* ⟵ *MBConv (w, exp = 6, out = 40, kernel = 5, stride = 1)*
11. *Repeat 3 times*
12. ⟶ *w* ⟵ *MBConv (w, exp = 6, out = 80, kernel = 5, stride = 1)*
13. *Repeat 3 times*
14. ⟶ *w* ⟵ *MBConv (w, exp = 6, out = 112, kernel = 5, stride = 1)*
15. *Repeat 4 times*
16. ⟶ *w* ⟵ *MBConv (w, exp = 6, out = 192, kernel = 5, stride = 1)*
17. ⟶ *w* ⟵ *MBConv (w, exp = 6, out = 320, kernel = 5, stride = 1*
18. ⟶ *w* ⟵ *Conv 1 × 1 (w, out_channels = 1280)*
19. ⟶ *w* ⟵ *BatchNorm (w)*
20. ⟶ *w* ⟵ *ReLU (w)*
21. ⟶ *w* ⟵ *Global_Average_Pooling (w)*
22. ⟶ *w* ⟵ *DenseLayer (w, 512)*
23. ⟶ *w* ⟵ *DenseLayer (w, 256)*
24. ⟶ *w* ⟵ *DenseLayer (w, 128)*
25. ⟶ Y_p ⟵ *Softmax (w) // Final prediction*
End

MBConv Block

1. ***MBConv(w, exp, out, kernel, stride)***:
2. ⟶ *w0* ⟵ *w*
3. ***If****exp > 1*:
4. ⟶ *w* ⟵ *Conv1 × 1(w, out_channels = exp × in_channels)*
5. ⟶ *w* ⟵ *BatchNorm(w)*
6. ⟶ *w* ⟵ *ReLU(w)*
7. ***End***
8. ⟶ *w* ⟵ *DepthwiseConv(w, kernel_size = kernel, stride = stride)*
9. ⟶ *w* ⟵ *BatchNorm(w)*
10. ⟶ *w* ⟵ *ReLU(w)*
11. ⟶ *w* ⟵ *Squeeze&Excitation(w)*
12. ⟶ *w* ⟵ *Conv1 × 1(w, out_channels = out)*
13. ⟶ *w* ⟵ *BatchNorm(w)*
14. ***If****stride = = 1 and shape(w) = = shape(w0)*:
15. ⟶ *w* ⟵ *w + w0 // Skip connection*
16. *return w*

Experimental setup

The Eff-ReLU-Net model is trained on 30 epochs using categorical cross-entropy loss, SGD optimizer with an initial learning rate of 0.01. Each of the 30 epochs involved shuffling the dataset to provide the model with a different group of samples, improving training. An early stopping technique with a patience value of 4 is also employed to avoid the model overfitting. The number of trainable parameters, average prediction time per sample, and peak GPU memory consumption of the proposed model are provided in Table 3. All experiments were executed on the system having an Intel(R) Core (TM) i7-6300U processor, 12 GB of RAM, and a 2GB NVIDIA m920 graphics card.

Table 3 Computational cost analysis of the proposed Eff-ReLU-Net

No. of parameters	5.3 million
Average Training Time	5–6 h
Average Inference Time	12.4 millisecond
Average GPU Usage	780 MB

Datasets

In our study, we utilized two publicly available wound datasets, Medetec [33] and AZH [34]. The Medetec dataset includes 12 classes of wound types with a total of 800 images. We specifically selected 4 main wound classes of this dataset, which are Diabetic Ulcers, Pressure Ulcers, Toe Ulcers and Venous Ulcers, comprising a total of 365 images. The dataset includes images with various wound characteristics, such as different sizes, colors, and stages of healing. The AZH dataset, obtained from a Milwaukee, Wisconsin, wound care center over two years, contains 730 images of four wound types (Diabetic, Pressure, Venous, and Surgical). Images are different in size, with widths varying from 320 to 700 pixels and heights from 240 to 525 pixels. The dataset is divided into train and test groups with a ratio of 75:25. Figure 2 shows some samples of wound images from both datasets.

Data augmentation

To enhance the classification performance of our proposed model, we tackle the limitations posed by small datasets using extensive data augmentation methods. These methods artificially increase the dataset size, guaranteeing more diverse representations for better model training. We used various augmentation techniques, including fixed angle rotations at 90°, 180°, and 270° to introduce rotational invariance, random rotation to get flexibility, and translation to obtain different perspectives. Moreover, elastic deformation is employed to simulate realistic distortions, and gamma correction is implied to regulate brightness and contrast levels. Figure 3 presents the original image along with the corresponding resized and augmented images. These augmentation approaches notably improve the size of the dataset. By leveraging these augmentation methods, we mitigate the limitations of utilized datasets, enhancing the generalization ability and robustness of our proposed model. The dataset statistics are provided in Table 4.

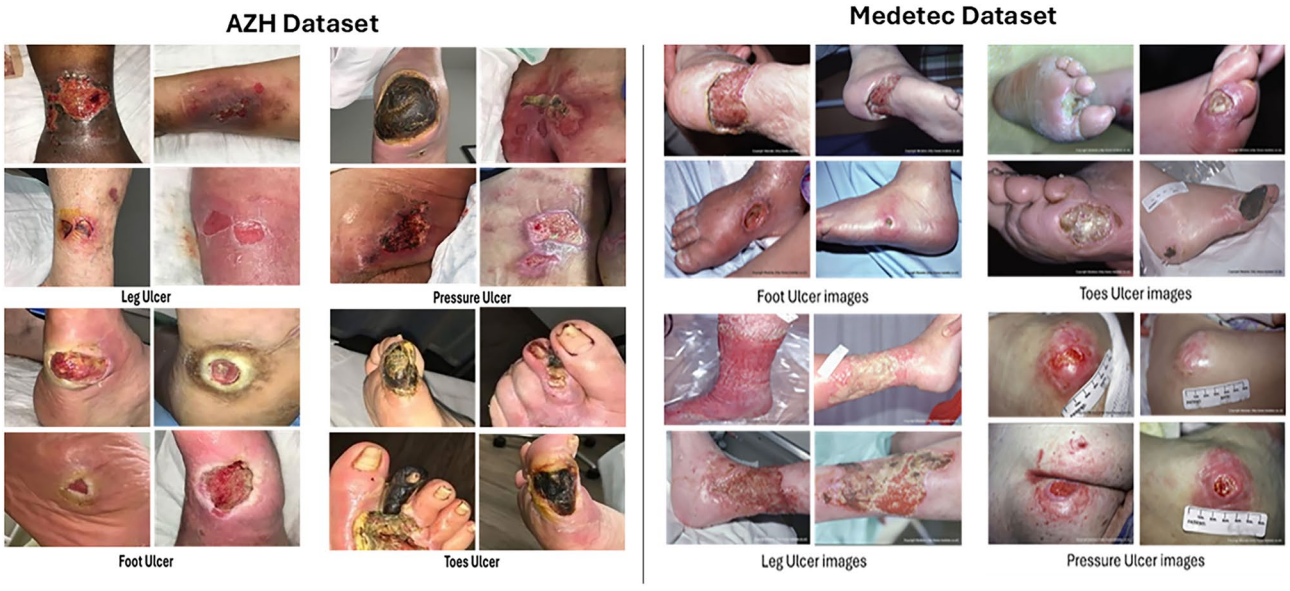


Fig. 2 Sample images for all four wound classes

Fig. 3 Samples of original images along with corresponding resized and augmented images

Evaluation metrics

To measure the suggested Eff-ReLU-Net, we have utilized the standard evaluation metrics of accuracy, recall, and precision.

Accuracy The ratio of accurately identified wound categories associated with the complete predictions, revealing the model's whole reliability. We calculated the accuracy as show in Eq. (2) [35].

Table 4 Dataset details

Dataset	Classes	Train Set		Test Set
		Before Aug	After Aug	
Medetec Dataset	Diabetic Ulcers	89	1270	30
	Pressure Ulcers	56	1081	19
	Toe Ulcers	26	891	9
	Venous Ulcers	102	1166	34
AZH Dataset	Diabetic Ulcers	128	1107	43
	Pressure Ulcers	131	1256	44
	Venous Ulcers	150	1650	50
	Surgical Ulcers	139	1304	46

$$Accuracy = \frac{TP + TN}{TN + FN + FP + TP} \quad (2)$$

Where the TP, TN, FP, and FN represent the true positives, true negatives, false positives, and false negatives, respectively.

Recall (Sensitivity) The ratio of accurately identified wounds of a specific class out of all actual classes, reflecting the model's ability to differentiate related cases. Recall is calculated as in Eq. (3) [35].

$$Recall = \frac{TP}{TP + FN} \times 100 \quad (3)$$

Precision The ratio of accurately detected wounds of a particular class out of all instances in which the model calculated for that class shows the accuracy of positive predictions. Precision is calculated as shown in Eq. (4) [35].

$$Precision = \frac{TP}{TP + FP} \times 100 \quad (4)$$

F1-score F1-score is a metric that balances the trade-off between precision and recall while differentiating between different wound types. It provided an accurate assessment in the presence of class imbalance and is calculated as shown in Eq. (5) [35].

$$F1 - Score = \frac{Precision \times Recall}{Precision + Recall} \times 100 \quad (5)$$

ROC curve The ROC curves plot the True Positive Rate against the False Positive Rate as the classification threshold varies. Each curve rises steeply toward the top-left corner, indicating strong discriminative power and the model's ability to distinguish different wound types.

Results and discussion

This section presents a detailed assessment of the suggested Eff-ReLU-Net model's performance for wound image classification. We have provided the details of the experimental setup, datasets, augmentation techniques, and different experiments designed to assess the competency of our method over the state-of-the-art (SOTA) methods and DL approaches. Furthermore, we also evaluated our method in a cross-corpora setting to measure the generalizability of the proposed method.

Table 5 Performance of Eff-ReLU-Net for Medetec and AZH datasets

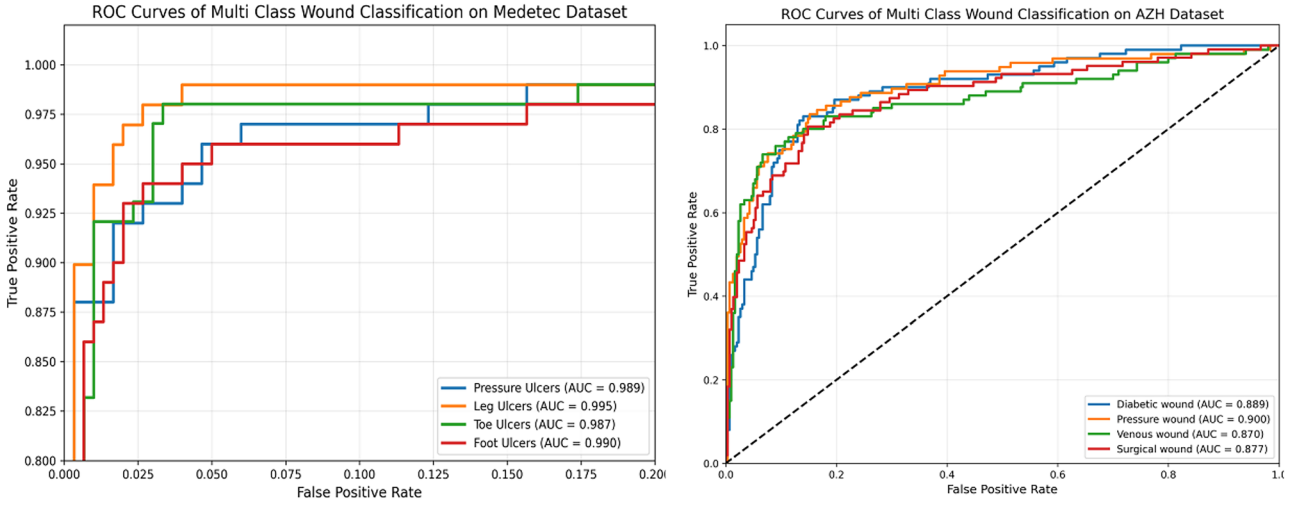
Dataset	Accuracy (%)	Recall (%)	Precision (%)	F1-Score (%)
Medetec Dataset	92.33	95.33	97.66	96.48
AZH Dataset	90.00	89.45	92.19	90.84

Performance evaluation

To evaluate the performance of the proposed Eff-ReLU-Net method for chronic wound classification, we conducted experiments utilizing the Medetec and AZH datasets. For this purpose, we trained our method to use the training collection of Medetec and AZH datasets separately and evaluated them on their corresponding testing collections with a split ratio mentioned in Sect. 4.2. The attained results for both datasets are reported in Table 3. From Table 5, it is depicted that the suggested approach performs well in classifying the wounds with an accuracy of 92.33%, a recall of 95.33%, a precision of 97.66%, and F1-score of 96.48% for the Medetec dataset. However, an accuracy of 90%, a recall of 89.45%, a precision of 92.19%, and F1-score of 90.84% is attained for the AZH dataset.

The ROC curves are also computed for AZH and Medetec datasets, as shown in Fig. 4. It can be observed that the Medetec dataset achieved a macro-average AUC of 99.03%, highlighting consistently strong discriminative performance across all wound classes. This high average indicates that the model performs well not just on one or two dominant classes, but across the entire set of classes, making it well-suited for handling class variability in real-world clinical data. Similarly, the AZH dataset attained a macro-average AUC of 88.4%, further demonstrating the model's reliable and balanced classification ability.

These results are attributed to the ability of the proposed method to obtain more distinctive and relevant features, obtained after introducing three fully connected dense layers and the incorporation of ReLU activation, which permits the flow of more relevant information. Combining the efficient scaling capability of the EfficientNetB0 model and ReLU activation allows the model to efficiently capture the distinct visual features in wound images, leading to improved wound recognition results on both datasets.



(a) ROC curve for Medetec dataset. **(b) ROC curve for AZH dataset.**

Fig. 4 ROC curves for wound datasets

Table 6 Comparison with other DL models utilizing the Medetec dataset

Model	Accuracy (%)	Sensitivity (%)	Precision (%)	F1-Score (%)
VGG-16	85.1	82.5	87.2	84.8
VGG-19	86.4	83.8	88.6	86.1
ResNet-50	88.03	85.77	91.01	88.31
Inception-V3	70.5	64.01	79.59	70.93
MobileNet	64.01	58.03	70.06	63.49
RepGhost.	82.3	85.1	82.8	83.94
RepVGG	85.3	80.1	84.5	82.24
MobileViT	83.7	86.2	88.3	87.24
Eff-ReLU-Net (Proposed Model)	92.33	95.33	97.66	96.48

Table 7 Comparison with other DL models utilizing the AZH dataset

Model	Accuracy (%)	Sensitivity (%)	Precision (%)	F1-Score (%)
VGG-16	76.06	75.40	79.10	77.20
VGG-19	70.23	73.02	72.41	72.71
ResNet-50	88.5	86.2	90.1	88.1
Inception-V3	78.3	75.8	82.4	78.9
MobileNet	74.01	68.03	80.06	73.49
RepGhost.	83.3	87.2	84.6	85.88
RepVGG	87.2	81.12	86.4	83.68
MobileViT	85.9	88.21	90.12	89.16
Eff-ReLU-Net (Proposed Model)	90.00	95.33	97.66	90.84

Comparison with other deep learning models

This experiment aims to analyze the effectiveness of the suggested Eff-ReLU-Net against the existing deep-learning models for wound classification. Specifically, we compared the proposed model performance with ResNet-50, VGG-19, Inception-V3, MobileNet-V2, RepGhost, RepVGG, and MobileViT utilizing the Medetec and AZH datasets. These models are nominated because of their huge employment in computer vision tasks and their proven effectiveness in feature extraction and classification. Each of these architectures represents a unique design, ranging from traditional DL approaches like ResNet-50 and Inception-V3 to lightweight architectures such as MobileNet-V2 and more advanced frameworks like RepGhost, RepVGG, and MobileViT, providing a comprehensive benchmark for evaluating the robustness and efficiency of the proposed model. Evaluating our model against these diverse approaches provided a comprehensive assessment of its ability to address the specific challenges of wound classification, like computing detailed wound features and handling variations in size and texture. The comparison in terms of accuracy, precision, recall, and F1-score is presented in Tables 6 and 7 for the Medetec and AZH datasets, respectively.

The results in Tables 6 and 7 demonstrate that the proposed Eff-ReLU-Net consistently outperformed all baseline models. It can be attributed to EfficientNetB0's unique compound scaling strategy, which optimally balances network depth, width, and resolution. VGG-16 and VGG-19 are renowned for their strong baselines with their deep, uniform structures, but lack the modern scaling techniques that improve efficiency. While ResNet-50 and Inception-V3, known for their deep architectures and strong feature extraction capabilities, lack the efficiency optimization found in EfficientNet, which limits their performance on specialized tasks like wound classification. MobileNet, although lightweight and efficient, compromises some accuracy due to its emphasis

on reducing computational load, which limits its ability to capture complex features in wound images. Furthermore, RepGhost and RepVGG both simplify their multi-branch blocks into single-path convolutions for faster inference but lack the compound-scaling needed for wound images, while MobileViT, which combines convolutions with self-attention, performs well on large datasets but adds unnecessary overhead on smaller datasets like Medetec and AZH. In comparison, our approach has overcome the issues of these methods by presenting a computationally simpler approach by employing the ReLU activation, which reduces gradient saturation issues and accelerates model convergence. Additionally, we incorporated average pooling and dropout layers to mitigate overfitting, which ensures better generalization across diverse wound categories. Furthermore, three fully connected dense layers are added to extract and pass distinctive features effectively to the SoftMax layer for accurate classification. So, Eff-ReLU-Net, with its efficient architecture, utilizes MBConvs and the ReLU activation function and adds dense layers, achieving superior performance by effectively balancing computational efficiency with enhanced feature extraction and attaining the best results on the Medetec and AZH datasets. Consequently, it demonstrates a balance between accuracy and efficiency that outperforms other leading architectures for wound classification. Highest accuracy, recall, precision, and F1-score on both wound datasets proved our proposed model the best choice for wound classification among other deep learning models.

Evaluation with State-of-the-Art methods

To show the superiority of our Eff-ReLU-Net over the existing wound classification approaches, we performed a comparative study of the suggested approach with SOTA methods. Particularly, we compared our method with the most recent approaches [17, 24, 26, 28, 36, 37] for both datasets. The comparison in terms of accuracy is shown in Table 8.

It is revealed from Table 8 that the proposed model attained the highest accuracy on both databases, with values of 92.33% and 90% for the Medetec and AZH datasets, respectively. Aldoulah et al. [24] employed a DL approach and attained an accuracy of 89.34%, while the work in [17] scored an accuracy of 91.90%. Moreover, the method in [28] used the feature concatenation-based approach and attained a score of 90%. Similarly [36], also utilized. feature concatenation approach implementing DenseNet and XceptionNet. Mousa et al. [37] utilized Xception model with gaussian mixture RNN and attained 88.81% accuracy. Our method secured the highest accuracy of 92.33%. Aldoulah et al. [28] employed EfficientNetB4, a larger and deeper variant of EfficientNet. While EfficientNetB4 provides strong feature extraction capabilities due to its depth, it also comes with increased computational costs. Our Eff-ReLU-Net, although smaller, utilizes a ReLU activation function that significantly improves computational efficiency while maintaining high classification accuracy, giving it an edge over the deeper EfficientNetB4. Rostami et al. [17] used an ensemble of Deep CNNs, which generally improves performance by combining multiple models to increase accuracy. However, ensemble methods tend to be computationally expensive and lead to diminishing returns when models are not sufficiently diverse. Our single well-optimized EfficientNetB0 model provided a more streamlined approach, balancing efficiency with accuracy without the added complexity of ensembling multiple networks. Alhababi et al. [28] implemented a feature concatenation-based method for classification. The feature concatenation-based method implies feature extraction employing numerous architectures that sum into a single feature vector and then classify into multiple wound classes using transfer learning and the transfer feature method. The method in [26] utilized numerous DL works with an attention mechanism, which resulted in model overfitting. Comparatively, we attained superior performance over both datasets due to the combination of EfficientNetB0's compound scaling strategy employing ReLU activation and the additional three fully connected dense layers. This balance allows our model to capture relevant features with high precision while maintaining computational efficiency, thus leading to the highest accuracy compared to the SOTA methods.

Table 8 Comparison with state-of-the-art methods on the medetect dataset

Year	Methods	Accuracy (%)	
		Medetec Dataset	AZH Dataset
2021	Ensemble CNN [18]	91.90	82.48
2023	Swish-ELU EfficientNet-B4 [19]	89.34	87.32
2024	Multi-modal Network [27]	83.00	78.00
2024	DenseNet121 + XceptionNet [35]	84.45	82.39
2025	DenseNet + MobileNet [26]	90.00	81.00
2025	Xception + GMRNN [36]	---	88.81
2025	**Eff-ReLU-Net (Proposed)**	**92.33**	**90.00**

Cross-Dataset evaluation

This experiment aimed to evaluate the generalization ability of our suggested Eff-ReLU-Net for chronic wound classification. To perform this task, we utilized the common classes among the Medetec and AZH datasets.

For this, we planned a cross-dataset assessment for the subsequent states: (a) trained the proposed model on the Medetec dataset and tested it over the AZH dataset, (b) trained Eff-ReLU-Net on the wound images of the AZH dataset and evaluated it utilizing the Medetec dataset.

The attained results are reported in Table 9, from where it can be seen that despite training on one dataset and testing on unseen samples of a completely different dataset, our method obtained reasonably good results for both cases, presenting the better generalizability of our architecture for the wound classification into relevant ulcers. The improved results in cross-dataset evaluation are due to the robustness of the proposed model in learning generalized features that effectively capture the complexities of different wound types. The incorporation of additional dense layers allowed the model to capture distinctive features across datasets, enhancing its adaptability. By focusing on common wound classes (diabetic, pressure, and venous wounds), the model achieved consistent performance despite variations in dataset characteristics, such as image quality and diversity, which indicates its potential for real-world clinical applications.

Table 9 Cross-dataset assessment results

Training Dataset	Testing Dataset	Accuracy (%)	Recall (%)	Precision (%)	F1-Score (%)
Medetec	AZH	68.33%	68.46%	68.66%	68.56
AZH	Medetec	70.15%	70.55%	70.33%	70.44

Table 10 Performance of EfficientNet variants on Medetec dataset

Model	Accuracy (%)	Recall (%)	Precision (%)	F1-Score (%)
EfficientNet-B1	87.6	86.3	88.9	87.6
EfficientNet-B2	88.5	87.2	89.8	88.5
EfficientNet-B3	89.8	88.5	91.1	89.8
EfficientNet-B4	90.4	89.2	91.7	90.4
EfficientNet-B5	90.9	89.7	92.2	90.9
EfficientNet-B6	91.0	89.8	92.3	89.9
EfficientNet-B7	91.1	89.9	92.4	91.1
Eff-ReLU-Net (proposed)	**92.33**	**97.66**	**95.33**	**96.48**

Table 11 Performance of EfficientNet variants on AZH dataset

Model	Accuracy (%)	Recall (%)	Precision (%)	F1-Score (%)
EfficientNet-B1	84.1	84.7	83.5	84.1
EfficientNet-B2	85.3	85.9	84.8	85.3
EfficientNet-B3	86.7	87.2	86.1	86.6
EfficientNet-B4	87.5	88.1	86.9	87.5
EfficientNet-B5	88.0	88.6	87.4	88.0
EfficientNet-B6	88.3	88.9	87.7	88.3
EfficientNet-B7	88.4	89.3	87.8	88.4
Eff-ReLU-Net (proposed)	**90.0**	**89.45**	**92.19**	**90.84**

Table 12 Impact of activation functions and number of fully connected layers

Model Variations	Accuracy (%) on Medetec Dataset	Accuracy (%) on AZH Dataset
EfficientNet B0 – Swish	86.5	80.2
EfficientNet B0 – ReLU	88.1	84.8
EfficientNet B0–2 FC Layers	88.1	85.9
EfficientNet B0–3 FC Layers	89.2	87.4
EfficientNet B0–4 FC Layers	88.7	87.2
EfficientNet B0 – ReLU – 3 FC Layers (proposed)	**92.33**	**90.0**

Ablation study

To examine the efficiency and efficacy of our proposed model, we conducted an extensive ablation study. We conducted two different types of experiments: (1) to analyze the performance of the proposed model among its variants and (2) to analyze the role of different activation functions and the number of fully connected layers in the multiclass wound classification task. The model variants are then trained and tested on the AZH and Medetec datasets for multi-class wound classification. The results of the ablation study presented in Tables 10, 11 and 12 confirm the robustness of the Eff-ReLU-Net model's current architecture for effectively classifying the different types of wounds.

Performance evaluation of model variants

In this subsection, we examine the performance of the proposed Eff-ReLU-Net in comparison with its architectural variants, primarily focusing on the underlying EfficientNet backbone choices (B1–B7) and our proposed modification based on EfficientNet-B0 with ReLU and additional dense layers.Tables 10 and 11 summarize the performance across both datasets.

The proposed Eff-ReLU-Net, which is based on EfficientNet-B0, enhanced with a ReLU activation function and three fully connected layers, consistently outperformed all other variants in the EfficientNet family. On the Medetec dataset, it achieved an accuracy of 92.33%, with an F1-score of 96.48%. On the AZH dataset, the model achieved 90.00% accuracy and a robust F1-score of 90.84%. EfficientNet-B7 is the second-top performer but at the cost of the highest number of parameters and FLOPS, while the EfficientNet-B1 model variant attained the lowest accuracy. Notably, performance gradually improved as the base model transitioned from B1 to B7, confirming the benefit of deeper EfficientNet variants. However, none surpassed our proposed model, suggesting that deeper networks alone are insufficient without the appropriate activation strategy and classification architecture.

Impact of activation functions and fully connected layers on accuracy

To further assess the performance of our architecture, we conducted targeted experiments on the EfficientNet-B0 model by varying two key components: the activation function (Swish vs. ReLU) and the number of fully connected (FC) layers (from 2 to 4). Table 12 presents a detailed comparative analysis based on accuracy utilizing the Medetec and AZH datasets.

From the results, it is observed that using the ReLU activation over Swish, particularly when combined with three FC layers, produces the best results for the multi-class wound classification. The proposed model EfficientNet-B0-ReLU-3FC attained the highest accuracy on both wound datasets and outperformed other model alternatives. Additionally, adding more than three dense layers (i.e., 4 FC) did not result in improved performance and sometimes led to slight overfitting. Therefore, three fully connected layers are the most appropriate choice to trade a balance between complexity and performance.

Statistical analysis

To analyze the statistical difference between the proposed Eff-ReLU-Net and other comparative deep learning models, we conducted statistical tests using ANOVA and Friedman approaches. For this analysis, we consider the Medetec dataset, and the utilized comparative deep learning models are: VGG-16, VGG-19, ResNet-50, Inception-V3, MobileNet, RepGhost, RepVGG, and MobileViT. We conducted these tests using the SciPy and Statsmodels packages in Python. The results for ANOVA and Friedman are summarized in Tables 13 and 14, respectively. The ANOVA test results in Table 13 reveal significant statistical differences between performance of the models, thus rejecting the null hypothesis that all models have the same mean performance for the wound classification. Furthermore, Friedman test results in Table 14 also confirm the performance difference among the models and attained the test statistics of 39.7867 with the p_value of 3.5106×10^(-6). Overall, both analysis tests reject the null hypothesis and provide statistical evidence that one model perform consistently better than other models.

Table 13 ANOVA test results

Types	Sum of Squares	Degrees of Freedom	F_Statistic	P_value
Treatments	3215.2684	8	14264.4410	1.2654×10^{-6}
Residuals	1.0143	36	--	--

Table 14 Friedman test results

Test	Value
Friedman test statistics	39.7867
p_value	3.5106×10
Null hypothesis	Reject

Discussion

Wound classification is a critical task in healthcare, particularly for managing chronic wounds. Accurate and efficient classification of these wounds is essential for timely and appropriate treatment, as the existing models often struggle with generalizing across diverse datasets and varying wound types. Therefore, developing a robust and efficient model that can handle diverse wound images with high accuracy is vital for improving clinical decision-making. The proposed Eff-ReLU-Net model leverages EfficientNetB0, enhanced with MBConvs and the ReLU activation function. Adopting ReLU activation in our EfficientNet architecture played a critical role in achieving faster convergence and improved generalization. By replacing the Swish function with ReLU, the model effectively addressed computational overhead and gradient saturation issues, resulting in more efficient learning. Additionally, the inclusion of dense layers enhanced the extraction of distinctive features, enabling the model to accurately classify wound types. These architectural improvements allowed the model to consistently adapt to diverse datasets, demonstrating its robustness and effectiveness in wound classification tasks. We assessed the model on the Medetec and AZH datasets, comparing it with other deep learning models and state-of-the-art approaches. Our model achieved accuracies of 92.33% and 90% over the Medetec and AZH datasets, outperforming several models, including ResNet-50, InceptionV3, MobileNet, RepGhost, RepVGG and MobileViT. Additionally, we compared our model to SOTA methods, where our approach demonstrated superior performance in wound classification. The ablation study was conducted to show the performance changes due to the different components and variants of the proposed models. The model was also evaluated in a cross-dataset setting and achieved a reasonable performance. These results emphasize the model's robustness and adaptability to different datasets, indicating its potential for widespread clinical use. The ability to achieve high accuracy on both datasets underlines the generalization capability of the model, which is crucial for real-world applications. Despite the promising outcomes, the proposed method needs to further improve the generalizability aspect. The datasets used, though effective, may not fully capture the diversity of real-world wound images, potentially affecting the model's generalization to unseen cases.

For future work, expanding the model's training with additional datasets, exploring advanced techniques such as transfer learning or self-supervised learning, and incorporating multimodal data can further improve its performance and assist in tackling diverse sample

distortions. These stages would make the model generalize better and provide more precise results across a wider range of clinical cases. Eff-ReLU-Net provides an effective and efficient solution for automated wound classification, demonstrating superior accuracy and adaptability across diverse datasets. By outperforming existing state-of-the-art models, the proposed approach highlights its potential for real-world clinical applications, particularly in enhancing diagnostic precision and efficiency.

Conclusion

This paper has presented the Eff-ReLU-Net method for automated chronic wound classification, evaluated by employing Medtect and AZH datasets. Experimental results, including the comparative analysis, demonstrate the proposed model's competence against the DL-based and SOTA wound classification approaches. The ablation study highlights the effectiveness of different components of the proposed Eff-ReLU-Net. Further, the cross-corpus evaluation is also conducted to check the generalization ability of the Eff-ReLU-Net. The experimental results underscore the effectiveness of the proposed model in accurately classifying chronic wounds, showcasing its robustness and potential for practical wound image classification. Moving forward, future research will focus on refining the architecture of the model by incorporating larger and more diverse wound datasets and conducting clinical trials to further enhance its accuracy, efficiency, and real-world applicability.

Acknowledgements
The authors extend their appreciation to the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) for funding this work through (grant number IMSIU-DDRSP2501).

Author contributions
Conceptualization Sifat Ullah, Ali Javed, Muteb Aljasem Methodology Sifat Ullah, Ali Javed, Muteb Aljasem, Abdul Khader Jilani Saudagar Software Sifat Ullah Validation Ali Javed, Muteb Aljasem Formal analysis Sifat Ullah, Ali Javed, Muteb Aljasem, Abdul Khader Jilani Saudagar Investigation Ali Javed Resources Ali Javed Data Curation Sifat Ullah, Muteb Aljasem Writing - Original Draft Sifat Ullah, Ali Javed, Muteb Aljasem, Abdul Khader Jilani Saudagar Writing - Review & Editing Sifat Ullah, Ali Javed, Muteb Aljasem, Abdul Khader Jilani Saudagar Visualization Sifat Ullah, Muteb Aljasem, Abdul Khader Jilani Saudagar Supervision Ali Javed Project administration Ali Javed, Abdul Khader Jilani Saudagar.

Funding
This work was supported and funded by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) (grant number IMSIU-DDRSP2501).

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 15 March 2025 / Accepted: 16 June 2025
Published online: 01 July 2025

References

1. Sen CK. Human wounds and its burden: an updated compendium of estimates. Adv Wound Care. 2019;8(2):39–48.
2. Sen CK, Gordillo GM, Roy S, Kirsner R, Lambert L, Hunt TK, Gottrup F, Gurtner GC, Longaker MT. Human skin wounds: a major and snowballing threat to public health and the economy. Wound Repair Regeneration. 2009;17(6):763–71.
3. International Diabetes Federation. IDF Diabetes Atlas. 10th ed. Brussels: International Diabetes Federation. 2021. Available from: https://diabetesatlas.org/idfawp/resource-files/2021/07/IDF_Atlas_10th_Edition_2021.pdf
4. Frykberg RG. Diabetic foot ulcers: pathogenesis and management. Am Family Phys. 2002;66(9):1655–63.
5. Chen L, Sun S, Gao Y, Ran X. Global mortality of diabetic foot ulcer: a systematic review and meta-analysis of observational studies. Diabetes Obes Metabolism. 2023;25(1):36–45.
6. Armstrong DG, Tan TW, Boulton AJ, Bus SA. Diabetic foot ulcers: a review. JAMA. 2023;330(1):62–75.
7. Chan KS, Lo ZJ. Wound assessment, imaging and monitoring systems in diabetic foot ulcers: A systematic review. Int Wound J. 2020;17(6):1909–23.
8. Gillespie BM, Walker R, Lin F, Roberts S, Nieuwenhoven P, Perry J, Birgan S, Gerraghy E, Probert R, Chaboyer W. Setting the surgical wound care agenda across two healthcare districts: A priority setting approach. Collegian. 2020;27(5):529–34.
9. Wei M, Yang D, Chen L, Wu L, Jiang Q, Niu N, Yang T. The prevalence and prevention of pressure ulcers: A multicenter study of nine nursing homes in Eastern China. J Tissue Viability. 2021;30(1):133–6.
10. Sugathapala RU, Latimer S, Balasuriya A, Chaboyer W, Thalib L, Gillespie BM. Prevalence and incidence of pressure injuries among older people living in nursing homes: A systematic review and meta-analysis. Int J Nurs Stud. 2023;148:104605.
11. Kocher G, Kumar G. Machine learning and deep learning methods for intrusion detection systems: recent developments and challenges. Soft Comput. 2021;25(15):9731–63.
12. Yadav DP, Sharma A, Singh M, Goyal A. Feature extraction based machine learning for human burn diagnosis from burn images. IEEE J Translational Eng Health Med. 2019;7:1–7.
13. Bansal N, Vidyarthi A. Multivariate Feature-based Analysis of the Diabetic Foot Ulcers Using Machine Learning Classifiers. InProceedings of the 2024 Sixteenth International Conference on Contemporary Computing. 2024;(pp. 527–534).
14. Veredas FJ, Luque-Baena RM, Martín-Santos FJ, Morilla-Herrera JC, Morente L. Wound image evaluation with machine learning. Neurocomputing. 2015;164:112–22.
15. Alevizos V, Arampidis N, Boja I, Papakostas GA. Reverse circular logarithmic LBP for diabetic foot ulcer detection. InMICCAI workshop on artificial intelligence over infrared images for medical applications 2024;(pp. 11–22). Cham: Springer Nature Switzerland.
16. Goyal M, Reeves ND, Rajbhandari S, Ahmad N, Wang C, Yap MH. Recognition of ischaemia and infection in diabetic foot ulcers: dataset and techniques. Comput Biol Med. 2020;117:103616.
17. Rostami B, Anisuzzaman DM, Wang C, Gopalakrishnan S, Niezgoda J, Yu Z. Multiclass wound image classification using an ensemble deep CNN-based classifier. Comput Biol Med. 2021;134:104536.
18. Hüsers J, Moelleken M, Richter ML, Przysucha M, Malihi L, Busch D, Götz NA, Heggemann J, Hafer G, Wiemeyer S, Babitsch B. An image based object recognition system for wound detection and classification of diabetic foot and venous leg ulcers. InChallenges of trustable AI and Added-Value on health 2022; (pp. 63–7).
19. Hüsers J, Hafer G, Heggemann J, Wiemeyer S, Przysucha M, Dissemond J, Moelleken M, Erfurt-Berge C, Hübner U. Automatic classification of diabetic foot ulcer images–a transfer-learning approach to detect wound maceration.

InInformatics and technology in clinical care and public health 2022;(pp. 301–4).
20. Aguirre Nilsson C, Velic M. Classification of ulcer images using convolutional neural networks.
21. Ahsan M, Naz S, Ahmad R, Ehsan H, Sikandar A. A deep learning approach for diabetic foot ulcer classification and recognition. Information. 2023;14(1):36.
22. Rostami B, Niezgoda J, Gopalakrishnan S, Yu Z. Multiclass burn wound image classification using deep convolutional neural networks. ArXiv Preprint ArXiv:2103.01361. 2021.
23. Sarp S, Kuzlu M, Wilson E, Cali U, Guler O. A highly transparent and explainable artificial intelligence tool for chronic wound classification: XAI-CWC. No. January. 2021;8:1–3.
24. Aldoulah ZA, Malik H, Molyet R. A novel fused multi-class deep learning approach for chronic wounds classification. Appl Sci. 2023;13(21):11630.
25. Kumar KA, Prasad AY, Metan J. A hybrid deep CNN-Cov-19-Res-Net transfer learning architype for an enhanced brain tumor detection and classification scheme in medical image processing. Biomed Signal Process Control. 2022;76:103631.
26. Patel Y, Shah T, Dhar MK, Zhang T, Niezgoda J, Gopalakrishnan S, Yu Z. Integrated image and location analysis for wound classification: a deep learning approach. Sci Rep. 2024;14(1):7043.
27. Abubakar A, Ugail H, Bukar AM. Can machine learning be used to discriminate between burns and pressure ulcer? InIntelligent Systems and Applications: Proceedings of the 2019 Intelligent Systems Conference (IntelliSys). 2020;(pp. 870–880).
28. Alhababi M, Auner G, Malik H, Aljasem M, Aldoulah Z. Unified wound diagnostic framework for wound segmentation and classification. Mach Learn Appl. 2025;19:100616.
29. Biswas S, Mostafiz R, Paul R, Mohi Uddin BK, Rahman KM MM, and, Shariful FNU. DFU_MultiNet: A deep neural network approach for detecting diabetic foot ulcers through multi-scale feature fusion using the DFU dataset. Intell Based Med. 2023;8:100128.
30. Biswas S, Mostafiz R, Paul BK, Uddin KM, Hadi MA, Khanom F. DFU_XAI: a deep learning-based approach to diabetic foot ulcer detection using feature explainability. Biomedical Mater Devices. 2024;2(2):1225–45.
31. Biswas S, Mostafiz R, Uddin MS, Paul BK. XAI-FusionNet: diabetic foot ulcer detection based on multi-scale feature fusion with explainable artificial intelligence. Heliyon. 2024;10(10).
32. Tan M, Le Q, Efficientnet. Rethinking model scaling for convolutional neural networks. InInternational conference on machine learning 2019;(pp. 6105–6114). PMLR.
33. Medetec Wound Database.: stock pictures of wounds, available at: https://www.medetec.co.uk/files/medetec-image-databases.html
34. AZH Dataset, available at. https://github.com/uwm-bigdata/wound-classification-using-images-and-locations
35. Vujović Ž. Classification model evaluation metrics. Int J Adv Comput Sci Appl. 2021;12(6):599–606.
36. Aljasem M, Alhababi M, Javed A, Aldoulah Z, Abouheaf M, Mayyas M, Al-Rousan W. Feature Concatenation-Based Deep Learning Method for Multiclass Wound Classification. In2024 International Conference on Electrical and Computer Engineering Researches (ICECER) 2024;(pp. 1–5).
37. Mousa R, Matbooe E, Khojasteh H, Bengari A, Vahediahmar M. Multi-modal wound classification using wound image and location by Xception and Gaussian Mixture Recurrent Neural Network (GMRNN). 2025.

Publisher's note