applied sciences

MDPI

Article

A Novel Fused Multi-Class Deep Learning Approach for Chronic Wounds Classification

Zaid A. Aldoulah [1,*], Hafiz Malik [2] and Richard Molyet [1]

[1] Department of Electrical Engineering and Computer Science, University of Toledo, Toledo, OH 43606, USA; richard.molyet@utoledo.edu
[2] Department of Electrical and Computing Engineering, University of Michigan-Dearborn, Dearborn, MI 48128, USA; hafiz@umich.edu
* Correspondence: zaid.aldoulah@gmail.com

Abstract: Chronic wounds affect the lives of millions of individuals globally, and due to substantial medical costs, treating chronic injuries is very challenging for the healthcare system. The classification of regular wound type is essential in wound care management and diagnosis since it can assist clinicians in deciding on the appropriate treatment method. Hence, an effective wound diagnostic tool would enable clinicians to classify the different types of chronic wounds in less time. The majority of the existing chronic wound classification methods are mainly focused on the binary classification of the wound types. A few approaches exist that classify chronic wounds into multiple classes, but these achieved lower performances for pressure and diabetic wound classification. Furthermore, cross-corpus evaluation is absent in chronic wound type classification, in order to better evaluate the efficacy of existing methods on real-time wound images. To address the limitations of the current studies, we propose a novel Swish-ELU EfficientNet-B4 (SEEN-B4) deep learning framework that can effectively identify and classify chronic wounds into multiple classes. Moreover, we also extend the existing Medetec and Advancing the Zenith of Healthcare (AZH) datasets to deal with the class imbalance problem of these datasets. Our proposed model is evaluated on publicly available AZH and Medetec datasets and their extended versions. Our experimental results indicate that the proposed SEEN-B4 model has attained an accuracy of 87.32%, 88.17%, 88%, and 89.34% on the AZH, Extended AZH, Medetec, and Extended Medetec datasets, respectively. We also show the effectiveness of our method against the existing state-of-the-art (SOTA) methods. Furthermore, we evaluated the proposed model for the cross-corpora scenario to demonstrate the model generalization aptitude, and interpret the model's result through explainable AI techniques. The experimental results show the proposed model's effectiveness for classifying chronic wound types.

Keywords: chronic wound; EfficientNet-B4; Medetec dataset; AZH dataset; wound classification

check for updates

Citation: Aldoulah, Z.A.; Malik, H.; Molyet, R. A Novel Fused Multi-Class Deep Learning Approach for Chronic Wounds Classification. *Appl. Sci.* **2023**, *13*, 11630. https://doi.org/10.3390/app132111630

Academic Editor: Thomas Lindner

Received: 29 September 2023
Revised: 20 October 2023
Accepted: 22 October 2023
Published: 24 October 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The term "wound" refers to the rupture of tissue's cellular structure at the place where the body is injured. Wounds can be categorized as acute and chronic, based on the extent of the physiological damage and the duration from the initial injury. In contrast to acute wounds that often recover within four weeks, chronic wounds resist healing within the expected time frame (i.e., six weeks) [1]. Chronic injuries are more severe, and they need proper treatment since such wounds are painful, sensitive, and can cause significant patient complications. According to an estimate, approximately 50% of hospitalized patients have injuries. However, chronic wounds affect 1% to 2% of the population in developed countries [2]. For instance, an estimated 8.2 million Americans suffer from chronic wounds [3], exerting considerable health issues, and adverse psychological and financial impacts. Chronic injuries mainly occur in the elderly population, and their management has a substantial economic impact on the healthcare system. It is also anticipated that chronic

wounds will continue to be an economic, social, and clinical challenge as the population grows. The worldwide chronic wound care market is expected to rise from $12.36 billion in 2022 to $19.52 billion in 2029, at a CAGR of 6.7% over the forecast period, according to the market research report published in March 2022 [4]. These statistics indicate the prevalence of chronic wounds, like a "silent epidemic", and their financial impact on the healthcare industry.

The most common types of chronic wounds include diabetic, pressure, venous, and arterial ulcers. These wound types prevent healing because of pathophysiological phenomena such as non-receptive epidermal or dermal cells to reparative stimuli, the existence of drug-resistant microbial biofilms, persistent infections, and prolonged inflammation [2]. Cardiovascular disease, diabetes, cancer, obesity, poor nutrition, and aging can also make it difficult for wounds to recover without specialized care. Sometimes, patients with injuries cannot gain access to specialized wound care and recommendations because of the shortage of wound specialists in primary healthcare sites, or due to pandemics like COVID-19 [5]. Furthermore, the correct wound class must be detected before the appropriate treatment and medication, which is manually performed by experts. Manual evaluation or the classification of a wound is time-consuming, and it requires a lot of experience and time to train a medical practitioner [5]. So, there is a need for the development of artificial intelligence (AI)-based diagnostic tools that can automatically classify chronic wounds and aid patients in remote or pandemic areas. AI methods can also assist practitioners in the reliable classification of chronic wounds, which ultimately helps in wound management and treatment, and saves time and cost.

The significance of improving healthcare techniques utilizing AI has been investigated for years. In the past few decades, artificial intelligence has demonstrated its effectiveness in medical imaging, including brain tumor detection, cancer screenings, COVID-19 detection, etc. Thus, it enhances precision medicine, improves medical screening, evaluates patient risk factors, and reduces the workload for clinicians. In the literature, researchers have introduced several machine learning (ML) and deep learning (DL) methods for wound classification, which include wound type classification, wound tissue classification [6,7], wound severity [8,9], and wound depth classification [10]. However, more research needs to be performed on the multi-class classification of chronic wound types. Most existing works [11–15] have focused on classifying the specific regular wound type, such as diabetic foot ulcer detection, venous ulcer detection, or pressure wound detection. Multiple chronic wound type classifications can allow the DL model to become generalized for different types of chronic wounds. So, a need still exists to develop an effective model and the performance enhancement of chronic wound classification. Therefore, we propose an effective and efficient end-to-end deep learning framework that can classify the multiple types of chronic wounds from wound images, including diabetic wounds, pressure ulcers, surgical wounds, and arterial and venous ulcers. More specifically, we have proposed a Swish-ELU EfficientNet-B4 (SEEN-B4) framework, which is a fusion of two EfficientNet-B4 models with different activation functions for extracting competent feature maps. The prediction module of our SEEN-B4 model comprises dense and dropout layers, where the L_2 regularization is applied to the second-last dense layer. Lastly, the extracted salient features are propagated to the softmax layer for accurate chronic wound classification. The major contributions of our research work include:

1. The proposal of a novel fused Swish-ELU EfficientNet-B4 model that concatenates Swish EfficientNet-B4 and ELU EfficientNet-B4 to improve the accuracy of classifying multiple chronic wounds.
2. An extension of the existing Medetec and Advancing the Zenith of Healthcare (AZH) datasets for 3-class wound classification, to reduce the class imbalance problem of these datasets.
3. An exhibition of the explainability power and the feature selection capability of the proposed framework.

4. Extensive experiments and cross-corpora evaluation for multi-class chronic wound classification to show the effectiveness and generalization aptitude of the proposed SEEN-B4.

The remainder of the article is organized as follows: Section 2 presents and discusses the literature review, and the proposed method is described in the Section 3. Details regarding the experimental setup and results, including datasets, performance evaluation, comparison with SOTA methods, explainability, cross-corpora evaluation, ablation study, and limitations, are provided in Section 4. Finally, Section 5 presents the conclusion.

2. Literature Review

This section discusses the existing approaches for the classification of chronic wound types. The existing approaches for wound type classification are mainly focused on binary classification; only a few studies have performed the multi-class classification. For instance, Refs. [16,17] introduced the CNN-based approaches for classifying diabetic foot ulcers. Abubakar et al. [18] suggested the customized ResNet50 for identifying healthy and burned skin, and achieved satisfactory results. In Ref. [19], an ensemble CNN network was introduced for the binary classification of infection and ischemia in diabetic wounds. Husers et al. [20] applied the transfer learning approach and presented a model based on Mobilenet-v1 to classify macerations in diabetic foot ulcers. The model attained an accuracy of 69% on the dataset having the 416 wound images collected from the wound care center of Christliches Klinikum Melle, Germany. Similarly, in Ref. [21], the Xception model with pre-trained weights was employed to classify diabetic and venous ulcers. The model was evaluated on a private wound image dataset (909 wound images), and 83% and 67% accuracy were achieved for cropped and full wound images, respectively.

Some studies [22–25] have also performed multi-class chronic wound classification. For instance, Rostami et al. [22] introduced an ensemble classifier consisting of two AlexNet classifiers for multi-class wound image classification. One classifier extracted the image-wise information, while the other processed the patch-level information using the sliding window technique. Finally, the combined output from the two classifiers was passed to the Multilayer Perceptron (MLP) to predict the class label. Patch generation and processing in patch-wise classifiers increase the ensemble model complexity regarding time and computational cost. For 3-class classification (diabetic, pressure, andvenous-arterial), average accuracies of 82.9% and 87.7% were obtained on the Medetec and AZH datasets, respectively. This method [22] has acceptable accuracy for classifying venous and surgical wounds, but it shows the lowest accuracy while classifying pressure and diabetic wounds. Overall, the model has good accuracy, but at the expense of time complexity. Salih et al. [23] suggested an approach that not only classified the wound images but also provided an explanation via the visualization of crucial areas where the model concentrated while classifying the wounds. The presented classifier employed pre-trained VGG-16, having to unfreeze the last three layers, along with the softmax layer. However, importance maps were generated by utilizing the LIME XAI tool and heatmap for explainability. The model was evaluated on their custom wound dataset, with the classes being named diabetic, pressure, surgical, and lymphovascular wound. This model [23] also showed lower results for diabetic and pressure wounds. In Ref. [24], a multi-modal network was presented that concatenated the two classifiers: (i) wound image classifier (WIC) and (ii) wound location classifier (WLC). The multi-modal classifier utilized the wound images and their corresponding locations for classifying the wound into one of the wound classes (diabetic, pressure, venous, or surgical). It is observed that utilizing only locations for wound classification yields a lower accuracy than when using only wound images, indicating that the data are not location-dependent. The CNN-based model demonstrated in [25] was evaluated on their private dataset with five wound types: deep wound, infected wound, arterial wound, venous wound, and pressure wound. The CNN network shared the same attention module-based backbone model, with task-specific branches. This model [25] has a classification accuracy of less than 75% for infected and deep wounds.

In wound diagnosis, wound severity classification can also aid clinicians in categorizing wound conditions. The transfer learning approach and stacked deep learning models were utilized to classify wound severity [26]. The dataset contained wound images of different types, classified into three classes: red, green, and yellow, by the wound specialist from the AZH center. A wound localizer based on the YOLO-v3 model was used to identify the region of interest, having a confidence score of greater than 97%. Multi-zoom learning networks were also presented in [26], each consisting of four models separately trained on the same image's zoom-out channels (Z0–Z3). The output of the four models was concatenated in the multi-zoom network for classification. The stacked models are incapable of outperforming the VGG-19 model. Moreover, multi-zoom learning has been identified as being ineffective in classifying the severity of wounds, due to its lower classification performance. The highest achieved classification accuracy was 68.49%, which needs to be improved. Other related information can be integrated to improve the classification results, such as the size and depth of the wound (which cannot be extracted from the image). Overall, the existing models for wound classification mostly perform binary classification. However, existing multi-class classification methods show an average performance and they need more explainability factors to increase the model's trustworthiness. Furthermore, the current wound classification approaches need to be assessed for cross-corpora scenarios, which is essential for analyzing model behavior for real-world scenarios.

3. Methodology

In this research, we have presented a regularized fused Swish-ELU EfficientNet-B4 (SEEN-B4) framework for classifying chronic wound types, including diabetic, pressure, venous, and surgical wounds. We have fused the two EfficientNet-B4 networks with different activation functions to effectively capture the complex and salient features present in the wound images. A detailed visual demonstration of the proposed approach is shown in Figure 1. The proposed approach comprises two main phases: (i) Pre-processing for augmenting wound dataset images and (ii) the SEEN-B4 framework for feature extraction and classification. The details regarding the proposed framework are provided in the subsequent subsections.

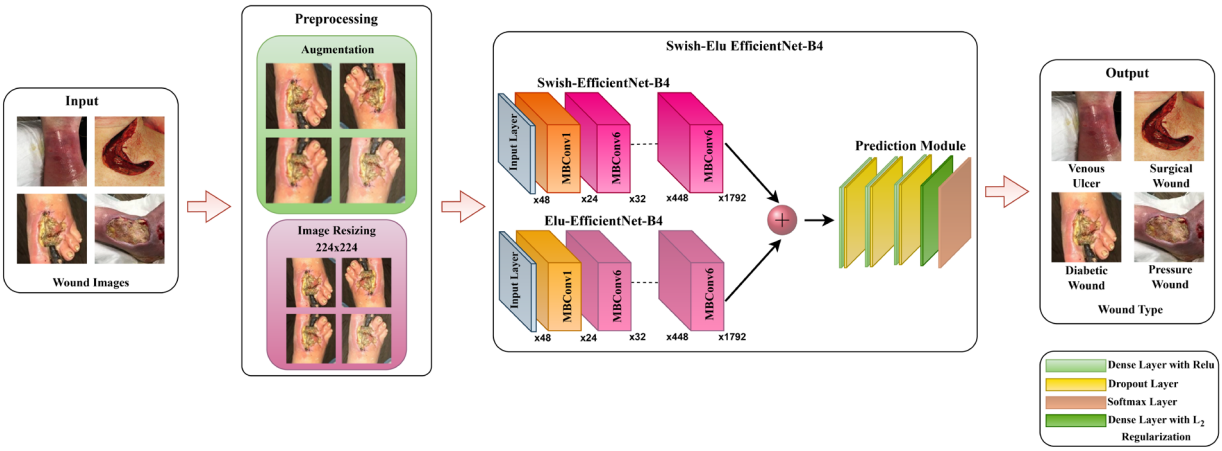


Figure 1. Data Processing and Analysis Pipeline of the Proposed Framework.

3.1. Pre-Processing

In the pre-processing stage, we augment the dataset, as the wound image datasets are small, and because DL-based models require many images for model training. We have applied 15 different augmentation techniques to increase the dataset size. The augmentation techniques used include Vertical Flip, Horizontal Flip, Vertical-Horizontal Flip, Sharp, Shift, Shear, Raise Saturation, Raise Hue, Raise Red, Raise Green, Raise Blue, Gaussian Noise, Gaussian Blur, Gamma Correction, Equalize Histogram. Some augmented image samples are shown in Figure 2, in the same order as mentioned above. After that, according to the

model requirement, the augmented images are resized to 224 × 224 resolution. The resized images are then fed to our SEEN-B4 model, which extracts the reliable set of features and classifies the input image into one of the wound classes.

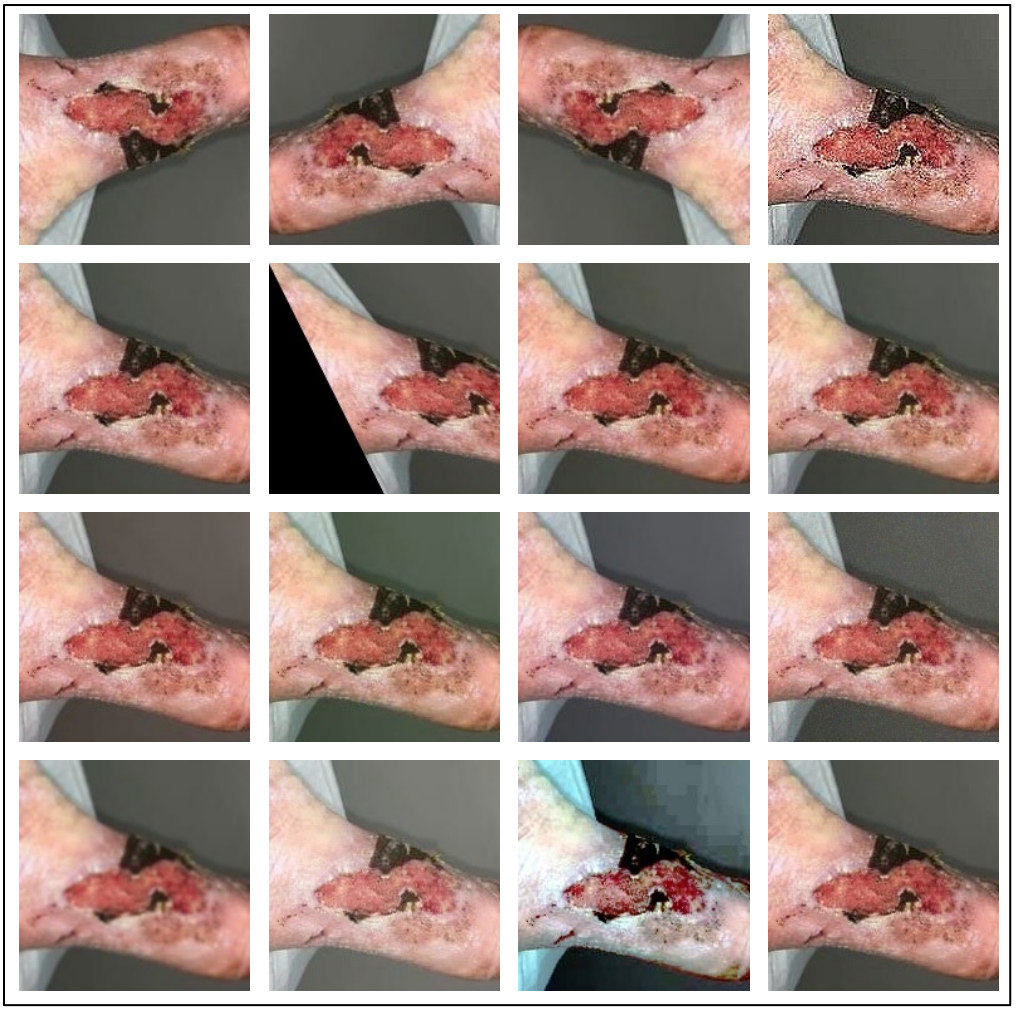


Figure 2. Augmented Wound Image Samples.

3.2. EfficientNet

To enhance the performance of earlier ConvNets, a combination of any two or of only one dimension (i.e., depth, width, and resolution of the networks) is scaled up. While scaling the model, adjusting all three network dimensions is essential for achieving a better efficiency and accuracy. EfficientNets introduced in 2019 [27] have utilized compound scaling to uniformly scale the network's depth, width, and resolution principally, and they have attained competent results. These models are optimized for parameters and FLOPs efficiency and they have different versions, from B0 to B7. EffiecientNet-B0 has fewer parameters and FLOPs, while the EfficientNet-B7 is computationally complex with the highest accuracy among the EfficientNet family. Inspired by the computational efficiency of EfficientNets and their reliable performance, we employed the EfficientNet-B4 network from the EfficientNet family in the proposed framework for classifying multiple wound types. We chose the EfficientNet-B4 based on the transfer learning experiment that we performed to analyze the different versions of EfficientNet while classifying the chronic wounds. The details of the experiment are provided in Section 4.5. EfficientNet-B4 has a series of MBConv1 and MBConv6 blocks. These blocks further have Expansion, Depth-wise, squeeze-and-excitation (SE), and output blocks. The expansion block contains Conv2D, Batch Norm, and Activation functions. However, the Depth-wise block

involves DepthConv2D, Batch Norm, and Activation function. In the SE block, the squeeze operation involves a global average pooling layer, and the excitation operation includes two fully connected layers, one with the ReLU and other employing a sigmoid activation function. After that, the obtained attention weights are multiplied element-wise with the input features. Both blocks (MBConv1 and MBConv6) are different in the output blocks. In MBConv1, the output block incorporates Conv2D and Batch Norm, while the output block of MBConv6 encompasses Conv2D, Batch Norm, and Dropout. Notably, the main building block of the EfficientNet-B4 is the MBConv having the SE optimization blocks. The SE blocks allow the model to learn to adjust the channel-wise feature maps via a gating mechanism. This enables the network to selectively emphasize more informative features by learning channel-wise weights and thus enhancing the accuracy.

3.3. Swish-ELU EfficientNet-B4

Our proposed framework consists of two EfficientNet-B4 models, which differ regarding the layer freezing and the utilized activation function. Specifically, we employed two models and fused them via concatenation to extract the more informative and salient features that would further help to accurately classify wound type. Each part of the proposed SEEN-B4 framework is described in the subsections.

3.3.1. Swish-EfficientNet-B4

The first model is the EfficientNet-B4 with a Swish activation function and layer freezing. We freeze the layers so the model learns generic high-level features from our dataset, stabilizes training, and prevents model overfitting. Additionally, the Swish activation function assists the model in improving gradient flow, generalization power, optimization, and prohibiting saturation. The Swish activation function allows for small negative values that are essential for learning complex patterns. The mathematical description of Swish is as follows:

$$S(z) = z \times sigmoid(\alpha z) \tag{1}$$

where z indicates the input to the Swish activation function, and α represents a learnable parameter. The architecture diagram of Swish-EfficientNet-B4 is shown in Figure 3.

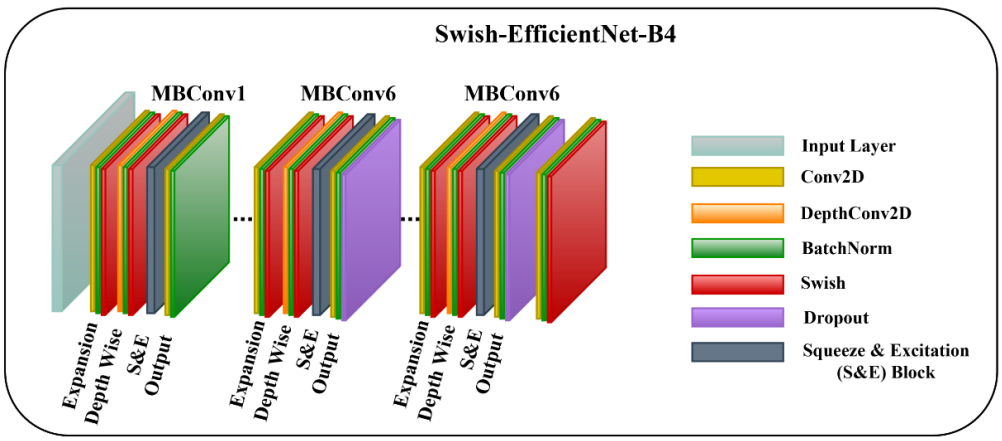


Figure 3. The Architecture of Proposed Swish-EfficientNet-B4 Framework.

3.3.2. ELU-EfficientNet-B4

In the other model, we employ the ELU activation function [28] in the EfficientNet-B4 network, and the model's layers are not frozen. Allowing the layers to be trainable helps the model to capture the low-level and task-specific features, which can lead to more powerful and relevant feature representations. Moreover, embedding the ELU activation function facilitates the model to learn discriminative features and capture complex patterns in the input data. ELU is a smooth and differentiable activation function that enables the model to

converge faster and is robust towards noisy inputs as it prevents extreme negative values. The mathematical representation of ELU is given as:

$$E(z) = \begin{cases} z & if\ z \geq 0 \\ \alpha \times (e^z - 1) & if\ z < 0 \end{cases} \quad (2)$$

where z is the input to the ELU activation function, and α is a hyperparameter that controls the extent to which an ELU saturates for negative information. Figure 4 represents the architecture diagram of ELU-EfficientNet-B4.

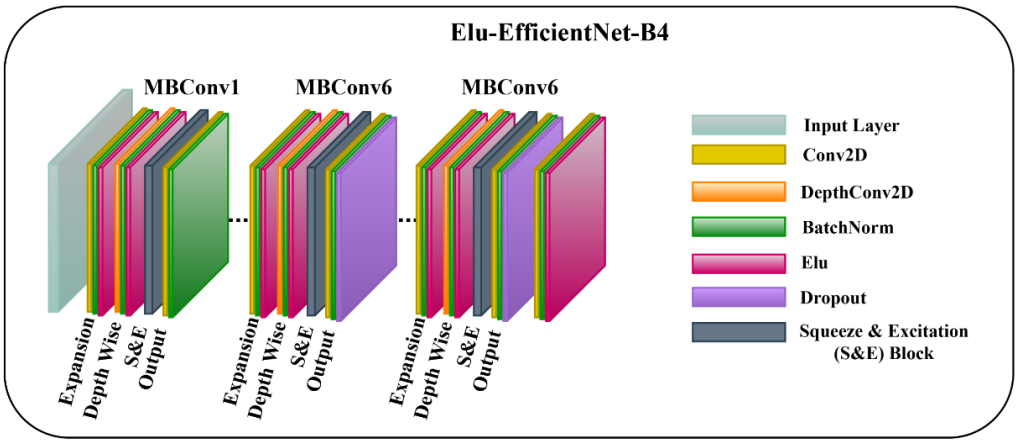


Figure 4. The Architecture of Proposed ELU-EfficientNet-B4 Framework.

3.3.3. EfficientNet-B4 Concatenation and Global Average Pooling Layer

After that, the two models (i.e., Swish-EfficientNet-B4 and ELU-EfficientNet-B4) are fused through the concatenation method. The features extracted from such a fused model are more informative and discriminative, which can help to further improve the overall classification accuracy of the model. The feature maps produced by the fusion of models are then passed to the global average pooling layer to reduce the spatial dimensions of the feature maps and to provide a compact representation that can capture more salient information. These feature representations are then propagated to the prediction module for classification.

3.3.4. Prediction Module

In the prediction module, we introduce three dense layers with ReLU activation, combined with the dropout layers. Incorporating dropout after each rampant layer aids with computing the competent features within the input data while minimizing the potential for overfitting. Then, a dense layer with L_2 regularization is introduced to prevent overfitting by penalizing large weights. The model learns more minor consequences, leading to a reduction in the complexity of the dense layer, and facilitating better generalization. The L_2 regularization added to the loss function of the dense layer can be represented as follows:

$$L_2 = \lambda \times \|M\|^2 \quad (3)$$

where λ represents the regularization parameter for controlling the regularization strength, and $\|M\|^2$ denotes the squared L_2 norm of the weight matrix M associated with the dense layer.

Finally, the extracted deep key features are passed to the last softmax layer to accomplish the classification task. The last layer of the proposed framework is the fully connected

layer with the softmax activation function that estimates the probabilities associated with each output unit, as follows:

$$\mathcal{S}(I_x) = \frac{e^{I_x}}{\sum_{n=1}^{K} e^{I_K}} \quad (4)$$

Here, K represents the number of classes, and I_x and I_K represent the input and output vectors, respectively.

4. Experimental Setup and Results

This section describes the datasets and performance evaluation measures used to evaluate the proposed framework. Also, the details of the extensive experiments and their results are discussed in the subsequent subsections.

4.1. Datasets

The effectiveness of the proposed SEEN-B4 model for chronic wound classification has been evaluated on publicly available Medetec and AZH datasets. In addition, we have also created extended Medetec and AZH datasets to evaluate the robustness of the proposed model. Table 1 presents the datasets details, and the description of the datasets is provided in the following sections.

Table 1. Datasets Details.

	Training		**Testing**		**Total**	
	Original	**After Augmentation**	**Original**	**After Augmentation**	**Original**	**After Augmentation**
			Medetec Dataset			
Diabetic	36	576	12	192	48	768
Pressure	130	2080	44	704	174	2784
Venous	100	1600	34	544	134	2144
			Extended Medetec Dataset			
Diabetic	130	2080	44	704	174	2784
Pressure	130	2080	44	704	174	2784
Venous	100	1600	34	544	134	2144
			AZH Dataset			
Background	75	1200	25	400	100	1300
Normal	75	1200	25	400	100	1300
Diabetic	139	2224	46	736	185	2960
Pressure	100	1600	34	544	134	2144
Venous	185	2960	62	992	247	3952
Surgical	122	1952	46	736	168	2688
			Extended AZH Dataset			
Diabetic	185	2960	62	992	247	3952
Pressure	185	2960	62	992	247	3952
Venous	185	2960	62	992	247	3952

4.1.1. Medetec Dataset

The Medetec wound dataset contains the free stock images of different types of wounds, including venous ulcers, pressure ulcers, burn injuries, and other miscellaneous wound types. Overall, this dataset has 14 wound types and 592 wound images. All of the images are in the jpeg format and have different sizes, with the width ranging from 358 to 560 pixels, and the height ranging from 371 to 560 pixels. For our experimentation, we have considered the three types of wounds, namely Diabetic, Pressure, and Venous.

4.1.2. Extended Medetec Dataset

The Extended Medetec dataset is the extended version of the original Medetec dataset. To reduce the class imbalance problem of the Medetec dataset, we have added more images in the Diabetic wound class. In the extended version, there are 174 images in the Diabetic wound class, while in the other two classes (i.e., Pressure and Venous), the number of images is the same as in the original Medetec dataset.

4.1.3. AZH Dataset

The AZH dataset was collected from the Wound care center in Milwaukee, Wisconsin, over two years. This dataset contains 730 wound images, having four different wound types, namely: Diabetic, Pressure, Venous, and Surgical wound. All the images are in the jpeg format, having varied sizes. The width of the images varies from 320 to 700 pixels, whereas the height varies from 240 to 525 pixels.

4.1.4. Extended AZH Dataset

As the name suggests, the extended version of the AZH dataset eliminates the class imbalance problem of the AZH dataset for 3-class classification. In the Extended AZH dataset, we have introduced the wound images in the Diabetic and Pressure wound classes. After the extension, the total number of images in each wound class, i.e., Diabetic and Pressure, becomes 247. In contrast, the Venous wound class has the same number of wound images (247) as in the AZH dataset.

4.2. Performance Evaluation Measures

To evaluate the performance of the proposed approach, we considered the standard evaluation metrics such as accuracy, precision, F1-score, and recall, as has also been adopted via the comparative approaches. Moreover, we have also used the confusion matrix analysis to report our method's wound classification performance.

Accuracy is calculated as the sum of the true positives (*TP*s) and true negatives (*TN*s) instances of each wound class, divided by the total number of cases. Overall accuracy is computed as:

$$Accuracy = \frac{\sum_{j=1}^{n}(TP_j + TN_j)}{\text{Total number of instances}} \tag{5}$$

Balanced Accuracy gives equal weight to each wound class while providing the model accuracy in case of an imbalanced dataset. It is computed as the average of the recall of each wound class. Balanced accuracy can be calculated as follows:

$$Balanced\ Accuracy = \frac{\sum_{j=1}^{n}(Recall_j)}{\text{n}} \tag{6}$$

Precision measures how accurately the proposed framework predicted each type of wound. Mean precision is calculated as:

$$Precision = \frac{\sum_{j=1}^{n}(TP_j)}{\sum_{j=1}^{n}(TP_j + FP_j)} \tag{7}$$

Recall represents the number of accurate optimistic predictions made of all the positive instances in the case of each type of wound in the dataset. Mean recall can be computed as follows:

$$Recall = \frac{\sum_{j=1}^{n}(TP_j)}{\sum_{j=1}^{n}(TP_j + FN_j)} \tag{8}$$

F1-Score represents the harmonic mean of the precision and recall of a model. For multi-class wound classification, it is computed as the arithmetic mean of the F1-Score of each wound class. The mean F1-Score can be calculated as follows:

$$F1 - Score = \frac{\sum_{j=1}^{n} \frac{2 \times Precision_j \times Recall_j}{Precision_j + Recall_j}}{n} \quad (9)$$

where, in Equations (5)–(9), n represents the total number of wound classes.

Confusion Matrix is considered as a visual evaluation tool for wound classification that summarizes the performance of the proposed model via enumerating true positives, false positives (*FP*s), true negatives, and false negatives (*FN*s). It aids in highlighting the wound classes that the proposed model misclassifies as other wound classes. This can help to interpret the model results in a better way, by analyzing the false rejection and acceptance scenarios.

4.3. Experimentation Protocols

The implementation of the proposed SEEN-B4 model is based on Keras TensorFlow. For each experiment, the dataset split ratio is 75:25, where 75% of the dataset is used for training the model, and 25% of the dataset is utilized for testing. The input images are resized to 224 × 224 resolution. The hyperparameters used to train the model are Batch size: 16, Epochs: 25, Adam optimizer with learning rate: 0.001, Categorial CrossEntropy loss, ReduceLROnPlateau with a factor: 0.5, patience: 2, and min learning rate: 0.000001. All the experiments are performed on a machine with the following specifications: 32 GB RAM, 12-core Processor at 3.70 GHz, 1 NVIDIA GeForce RTX 3090 GPU, and Windows 10 Pro.

4.4. Performance Evaluation of the Proposed Model

We conducted rigorous experiments to demonstrate the effectiveness of our proposed SEEN-B4 model for multi-class wound classification on the AZH dataset, Medetec dataset, and their extended versions. The experimental protocols, the results, and a discussion of the experiments are presented in the subsequent sections.

4.4.1. Evaluation on the Medetec Dataset

To examine the efficacy of the proposed model for the Medetec dataset, we conducted a 3-class classification experiment with the wound classes Diabetic, Pressure, and Venous. Each class of the dataset is split into training and testing sets. Then, we applied augmentations to the training samples to increase the number of images in the training set. Among the training samples, 80% of the data were used for training while 20% was used for validation. The trained model was then evaluated on the unseen examples in the testing set. The model achieved an accuracy, precision, recall, and F1-score of 88%, 84%, 79%, and 81.42%, respectively, on the Medetec dataset.

Moreover, we also generated the confusion matrix of our method on the Medetec dataset, and results are presented in Figure 5. From the confusion matrix, it can be observed that our SEEN-B4 model classified the Pressure wounds with the highest accuracy, while the Diabetic wounds were not classified accurately. More specifically, we observed three false positives and six false negative instances for the Diabetic wound. Moreover, half of the Diabetic wounds were misclassified as Pressure wounds. This might be because the Diabetic wound images are significantly fewer in number in the Medetec dataset, which could make it difficult for the model to learn the Diabetic wound features. The model is more biased toward Pressure wounds since the *FN* is 1 for Pressure wounds, and because the number of images in the Pressure wounds class is the highest. Our model has attained a reasonably good accuracy on the Medetec dataset, considering the vast class imbalance problem and the inconsistent standards of the acquired images.

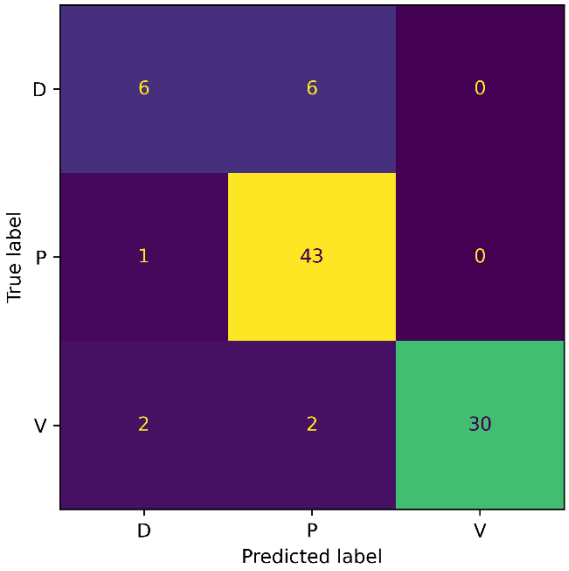


Figure 5. Confusion Matrix for Medetec.

4.4.2. Evaluation on the Extended Medetec Dataset

To mitigate the impacts of the enormous class imbalance in the Medetec dataset, we designed an experiment where we evaluated our proposed model on the Extended Medetec dataset. The experimental protocols are the same as in Section 4.4.1, and the details of the training and testing samples in each class are provided in Table 1. Our proposed SEEN-B4 model obtained an accuracy of 89.34%, and a precision, recall, and F1-score of 89.67% on the Extended Medetec dataset. The confusion matrix for this experiment is shown in Figure 6.

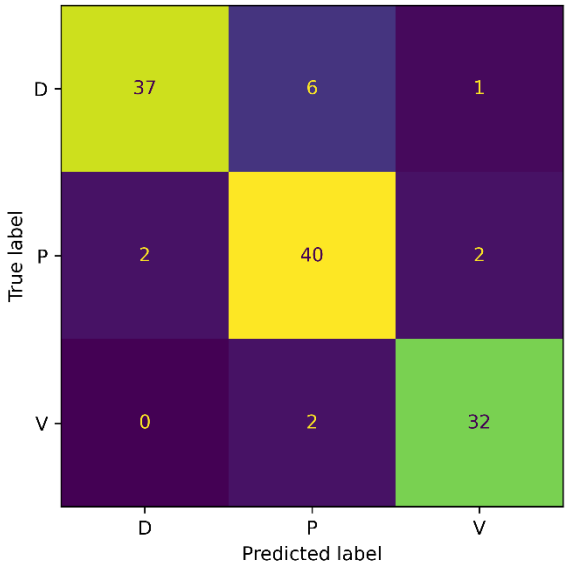


Figure 6. Confusion Matrix for Extended Medetec.

The confusion matrix shows that the accuracy for the Diabetic wound class increased from 50% to 84% with the increase in the number of images. Moreover, the classification accuracy for the Venous wound increased, while it decreased for the Pressure wounds class, compared to the results on the Medetec dataset. More precisely, we noticed *FP*s of 2, 8, and 3, and *FN*s of 7, 4, and 2, for the Diabetic, Pressure, and Venous wounds, respectively. We also observed that Diabetic wounds are mainly classified as Pressure wounds because these wounds are similar in shape and texture to the Pressure wounds on the foot area. In general, reducing the class imbalance problem of the dataset has enabled the model

to become less biased toward the Pressure wound class and has improved the proposed model's overall classification accuracy. These results signify the augmentation process's importance in addressing the class imbalance issue of original datasets in our method.

4.4.3. Evaluation on the AZH Dataset

To assess the performance of our model on the AZH dataset, we designed two experiments with varying numbers of wound classes. The first experiment is performed on the three classes of wounds, i.e., Diabetic, Pressure, and Venous ulcers, whereas the second experiment is conducted on six classes. Among the six classes, four are the wound classes, such as Diabetic, Pressure, Venous, and Surgical, while the other two are the non-wound classes, namely, Background and Normal. The model is trained on the augmented training samples for each experiment and is evaluated on the testing set samples.

For the 3-class classification experiment, our model attained an accuracy of 87.32%, a precision of 88%, a recall of 86%, and an F1-score of 86.98%. Whereas for the 6-class experiment, the proposed model achieved 83.19% accuracy, 84.34% precision, 83.67% recall, and 84% F1-score. It is noticed that as the number of classes increases, an accuracy drop of 4% occurs. The possible reason for this is that with an increasing number of classes, the chance of misclassification increases, eventually resulting in the model's performance drop. Additionally, the number of model parameters increases with the number of classes, making it more challenging for the network to train all of the parameters to the same degree, as with the smaller types.

Figures 7 and 8 show the confusion matrix results for both experiments: 3-class, and 6-class classification, respectively. Among the wound classes, it can be observed from the confusion matrix that the model has achieved the highest accuracy for Venous wounds.

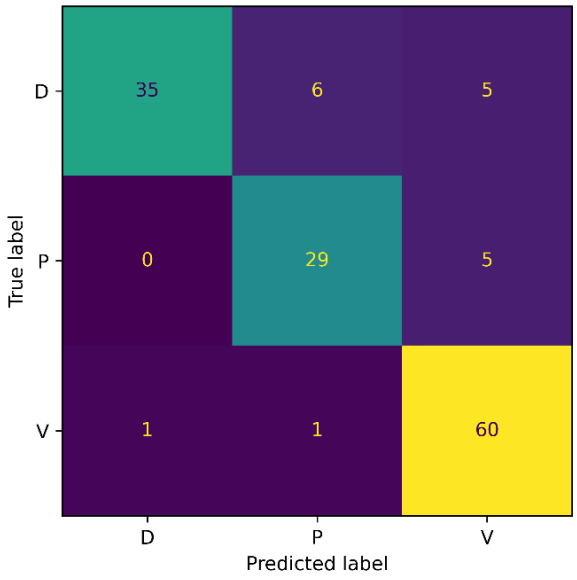


Figure 7. Confusion Matrix for AZH 3-class.

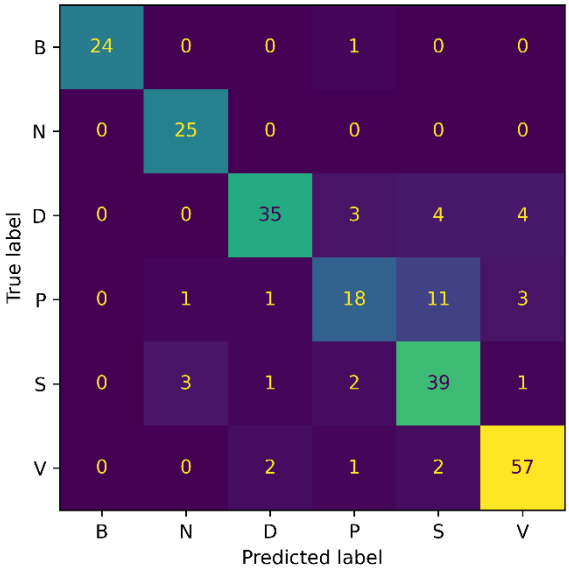


Figure 8. Confusion Matrix for AZH 6-class.

For the 3-class classification, the lowest accuracy for the Diabetic wounds can be observed, because the Diabetic wound class has the highest false negative and lowest false positive rates, compared to the Venous and Pressure wounds. However, in the case of the 6-class classification, the model attained the lowest accuracy for the Pressure wounds since the false negatives of 16 could be observed. Moreover, it can also be noticed that Pressure wounds are mostly classified as Surgical wounds. This can be due to the reason that Pressure wounds have similarities to some Surgical wounds, in terms of shape, color, size, and texture. Additionally, non-wound classes are accurately identified by the model, with classification accuracies of 96% and 100%. Overall, the classification ability of the proposed SEEN-B4 for the AZH dataset is relatively reasonable. However, the results presented in the confusion matrices reveal that the Pressure and Diabetic wounds are the most difficult to classify.

4.4.4. Evaluation on the Extended AZH Dataset

We experimented on the Extended AZH dataset to check the model performance for the balanced dataset. For this purpose, we performed a 3-class classification experiment. The experimental protocols are the same as for the 3-class classification experiment in Section 4.4.3, and the details of the training and testing samples in each class are provided in Table 1. The model achieved 88.17% accuracy and 88.34% recall, precision, and F1-score. The confusion matrix is illustrated in Figure 9.

The classification accuracy for the Diabetic and Pressure wound classes improved, as observed from the confusion matrix. However, compared to the results from the AZH dataset, a decrease in the accuracy of the Venous wound class was noticed. We observed the *FP*s of 10 and *FN*s of 4 for Venous wounds. The Diabetic wound class has 8 *FN* instances that are misclassified as Pressure and Venous wounds. This is because Diabetic wounds are identical to Pressure wounds in terms of shape, while they are similar in texture to Venous wounds. Generally, the overall classification accuracy of the model has improved because of the balanced dataset.

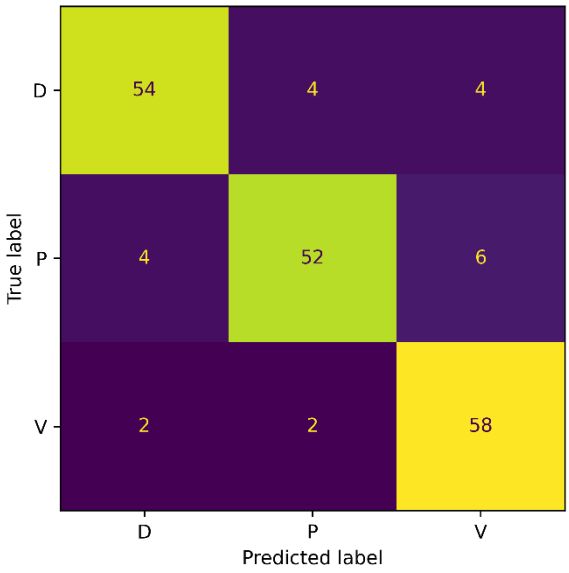


Figure 9. Confusion Matrix for Extended AZH.

Some of the wound images that are misclassified by our model are presented in Figure 10. Overall, it is observed that the shapes and textures of some diabetic wounds are similar to the pressure wounds on the foot, such as having a fluid texture and some yellow slough around the wound. Also, some of the diabetic wounds have a texture that is similar to the venous wounds; for instance, a dry crust. Therefore, diabetic wounds are misclassified as pressure and venous wounds, and vice versa. Pressure wounds having scab and dry crust are misclassified as venous wounds, due to the similarity in structure to venous wounds. Likewise, the venous wound with liquid and yellow slough on it is wrongly classified as a pressure wound.

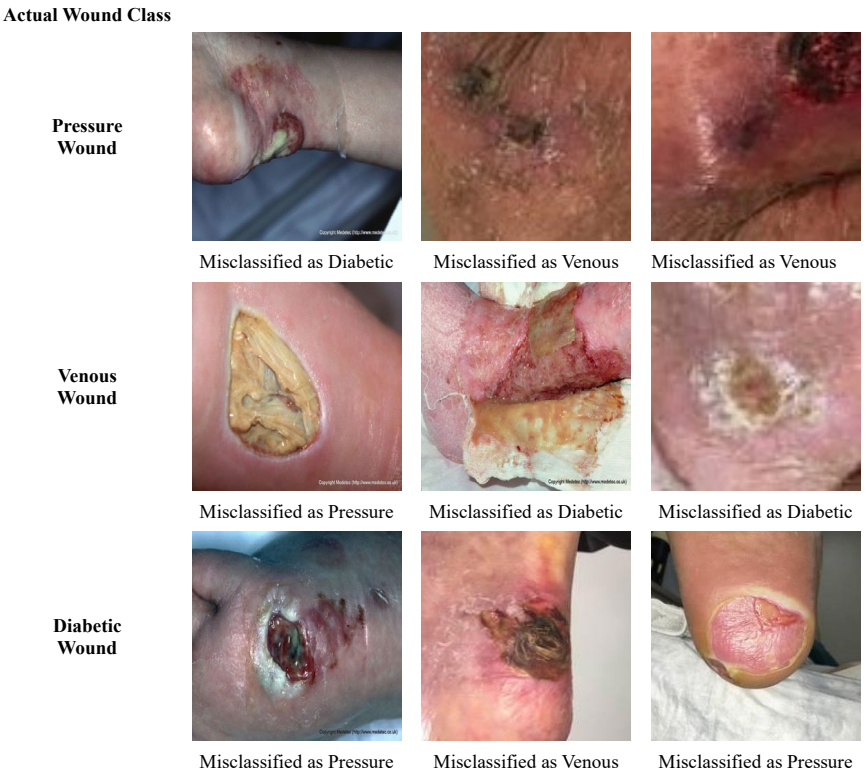


Figure 10. Misclassified Samples.

The summarized performance of our proposed SEEN-B4 model, in terms of accuracy, balanced accuracy, precision, recall and F1-score, on the Medetec, Extended Medetec, AZH, and Extended AZH datasets, is presented in Table 2. In Figure 11, we also presented the area under the receiver operating characteristic (ROC) curves. In the case of both the 3-class and 6-class classifications, Figure 11 shows that we have a high area under the curve (greater than 92%) for each wound class, except for the Diabetic wound class of the Medetec dataset. Since the Diabetic class has very few samples compared to other two classes in the Medetec dataset; therefore, when we increased the diabetic samples in the Extended Medetec dataset, the area under the curve for Diabetic class became 95%. The high area under the ROC curve indicates the ability of our model to effectively discriminate and classify the multiple types of wounds.

Table 2. Performance of the Proposed SEEN-B4 Model.

Dataset	Accuracy (%)	Balanced Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Medetec, 3-class	88.00	79.00	84.00	79.00	81.42
Extended Medetec, 3-class	89.34	89.71	89.67	89.67	89.67
AZH, 3-class	87.32	86.05	88.00	86.00	86.98
Extended AZH, 3-class	88.17	88.17	88.34	88.34	88.34
AZH, 6-class	83.19	83.62	84.34	83.67	84.00

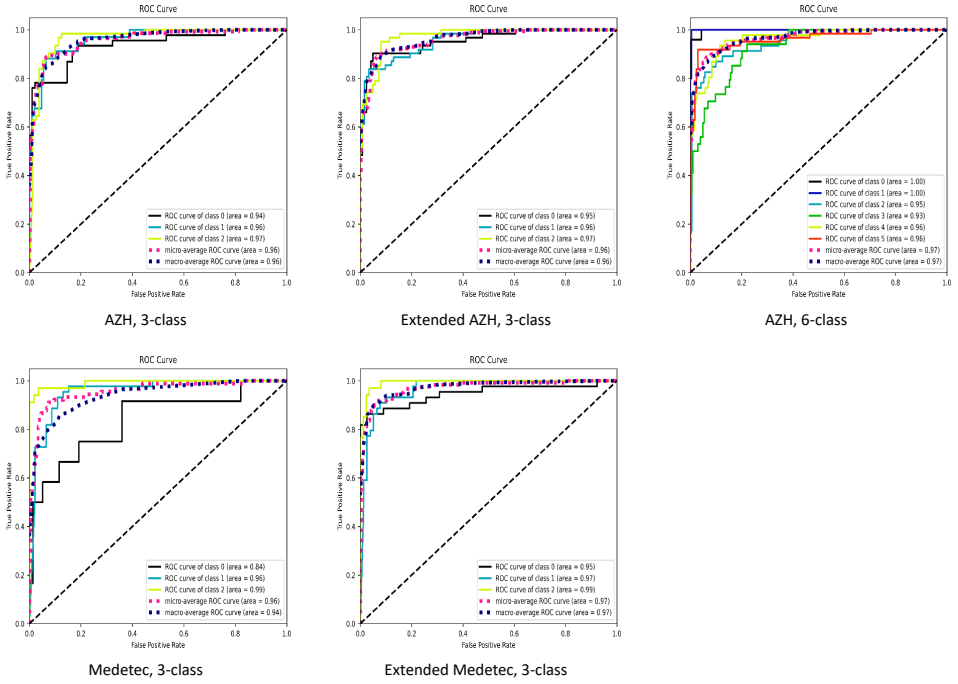


Figure 11. ROC curves. For 3-class classification, class 0: Diabetic, class 1: Pressure, class 2: Venous ulcer. For 6-class classification, class 0: Background, class 1: Normal, class 2: Diabetic, class 3: Pressure, class 4: Surgical, class 5: Venous ulcer.

The AZH and Medetec datasets have a class imbalance, while the class imbalance is diminished in the Extended AZH and Extended Medetec datasets. These datasets contain images that can have inconsistent standards because of the image acquisition being performed by multiple personnel, having distinct imaging equipment, differences in the angle and distance of the capture, and variations in the imaging content such as light intensity, color mode, etc. Consequently, this makes the classification task a challenging one. However, the experimental results demonstrate that our proposed framework can distinguish the different types of chronic wounds, such as Diabetic, Pressure, Venous, and Surgical wounds.

4.5. Comparison of Different Versions of EfficientNet

We designed this experiment to choose the EfficientNet version for our proposed framework that classifies chronic wounds more accurately. For this purpose, we fine-tuned the EfficientNets (B0–B7) to classify injuries. We trained the different EfficientNets on the augmented training set of both datasets (AZH and Medetec) for 3-class classification, and then we evaluated the trained model on the testing set. The results of the experiment are reported in Table 3.

Table 3. Results of EfficientNets on AZH Dataset.

EfficientNets	Medetec Dataset			AZH Dataset			
	Accuracy (%)	Precision (%)	Recall (%)	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
EfficientNet-B0	81.12	74	73	80.28	80	77	78.47
EfficientNet-B1	84.45	79	82	75.35	76	74	74.99
EfficientNet-B2	77.76	75	70	75.35	76	72	73.94
EfficientNet-B3	85.00	83	78	71.12	70	70	70.00
EfficientNet-B4	86.67	82	80	81.00	80	82	80.99
EfficientNet-B5	84.45	85	76	76.00	75	77	77.00
EfficientNet-B6	82.23	77	73	76.76	77	74	75.47
EfficientNet-B7	82.23	82	72	76.76	78	74	75.96

EfficientNet B5-B7 are larger models with more depth and several parameters. These models are computationally complex and lead to the model overfitting problem because of the smaller size of the wound datasets. From the results in Table 3, it can be interpreted that the EfficientNet-B4 classifies the wound types with the highest accuracy. So, based on these results, we utilized EfficientNet-B4 in our proposed framework, SEEN-B4.

4.6. Comparative Study

We compared the proposed SEEN-B4 framework with the existing contemporary wound classification methods [22,24] to assess its effectiveness over the comparative approaches for wound classification. Specifically, the performance of SEEN-B4 on the AZH dataset is compared with the wound image classifier (WIC) [24] and wound multimodality classifier (WMC) [24]. Furthermore, for the Medetec dataset, we compared the results with the ensemble classifier [22], WIC [24], and WMC [24]. In WMC, the outputs of two models (image and location classifiers) were concatenated to produce the final prediction, while WIC only considered the image of the wound for classification. A VGG-based backbone network was employed in both classifiers. VGG consists of stacked convolutional layers with relatively more parameters than EfficientNet, which may affect its performance. In contrast, EfficientNet efficiently and uniformly scales the model's width, depth, and resolution using compound scaling, allowing it to achieve a more excellent performance with fewer parameters than the VGG. A comparison of the methods for 6-class classification on the AZH dataset, and for 3-class classification on the AZH and Medetec datasets, is shown in Table 4.

Table 4. Comparison with Existing Methods on AZH and Medetec Datasets.

Datasets	Models	Accuracy (%)
AZH, 3-class	WIC [24]	76.06
	WMC [24]	84.51
	SEEN-B4 (Ours)	87.32
AZH, 6-class	WIC [24]	75.64
	WMC [24]	82.48
	SEEN-B4 (Ours)	83.19
Medetec, 3-class	Ensemble Classifier [22]	82.90
	WIC [24]	82.22
	WMC [24]	86.67
	SEEN-B4 (Ours)	88.00

It is noticeable that the proposed SEEN-B4 model provided superior classification accuracy over the existing method, for both the 3-class and 6-class classifications. More precisely, our SEEN-B4 attained the highest accuracy on the AZH and Medetec datasets, while the WMC [24] is the second-best performing model. Additionally, for 6-class classification, our model obtained an average accuracy gain of 4% from the existing method. At the same time, an average accuracy gain of 7% and 4% was achieved for 3-class classification on the AZH and Medetec datasets, respectively. Overall, the proposed framework outperforms the contemporary wound classification methods for the multi-class classification on both datasets, demonstrating its capability to classify multiple wound types.

4.7. Explainability

Due to the patient's health concerns, the performance, trustworthiness, and explainability of the model classifying the medical images is essential. The black-box nature of the deep learning classifiers makes them less adaptable to clinicians/surgeons, since they cannot trust the obscure cause on which these models generate predictions. So, it is necessary to interpret the predictions of such models by visualizing and inspecting the outputs of the model's layers. Therefore, we conducted an analysis to assess the explainability power of the proposed framework for wound classification. For this purpose, we applied the gradient-weighted class activation mapping (Grad-CAM) [29] to the last convolution layer of the proposed model to interpret the model's predicted output via the generation of heatmaps. It is anticipated that the previous convolution layer extracts abstract and salient deep features that influence the final decision. So, the generated heatmaps highlight the critical regions of the wound image sample that endorse the model predictions, since the model focused on those regions to make the final predictions. Such interpretations can enhance the confidence of clinical experts with regard to wound classifier decisions, using a visual explanation. The original wound images and their corresponding heatmaps generated through the Grad-CAM are shown in Figure 12 for the different types of wounds, including Pressure, Diabetic, and Venous wounds. The visual regions with bright colors (red, orange, and yellow) represent more focused pixels, while the blue and green colored areas are less focused by the model.

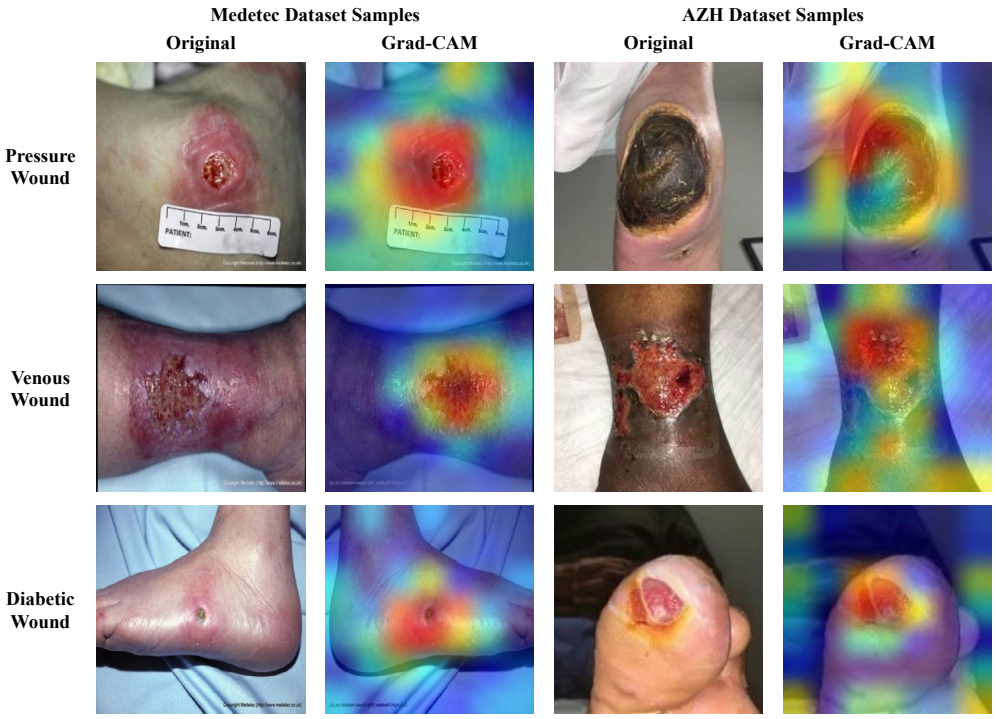


Figure 12. Heatmaps Visualization of Wound Images.

From Figure 12, it can be observed that the SEEN-B4 model identifies the precise location of the different types of wounds as it concentrates on the area where the wound exists in the image. Specifically, our model focuses more accurately on the reddish wound areas, and it is sometimes unable to precisely focus on the area where the eschar or dry crust is present. Overall, our model focuses on the more profound and darker damaged site, wound scar, and surrounding wound tissues. This reveals our fused framework's effective feature selection and learning capability while classifying the multiple types of wounds.

4.8. Ablation Study

We conducted an ablation study to analyze the effects of different activation functions on the accuracy of our SEEN-B4 model. For this, the first EfficientNet-B4 model contains a Swish activation function with layer freezing; however, we employ different activation functions in our second EfficientNet-B4 model. Specifically, we conducted experiments with Swish, SeLU, ReLU, and ELU activation functions for the second EfficientNet-B4 model in our fused method. The models are then trained and tested on the AZH dataset for 3-class classification. The results of the ablation study are presented in Table 5. The results indicate that our proposed model with the Swish-ELU activation function combination outperforms all the other variants and achieved an overall accuracy gain of 4.69%. SEEN-B4 with Swish-Swish activation is the second-best performer, while the model with the Swish-SeLU variant attained the lowest accuracy. So, Table 5 indicates that our proposed model with Swish-ELU activation function can effectively classify the different types of wounds. ELU is the non-saturated activation function that smooths slowly, is robust towards noise, and allows the activations' mean to become closer to the natural gradient. Thus, it generates accurate outputs and improves the model's performance.

Table 5. Ablation Study.

Dataset	Accuracy (%)	Balanced Accuracy (%)
Swish-SeLU	80.28	77.64
Swish-ReLU	81.69	78.58
Swish-Swish	85.91	84.39
Swish-ELU (Our Method)	87.32	86.05

4.9. Cross-Corpora Evaluation

We conducted a cross-corpora experiment to check the generalizability aptitude of the proposed SEEN-B4 model for multi-class wound classification. It has been noted that the existing works on wound classification have not investigated the aspect of the cross-corpora evaluation of the model. So, we performed a cross-corpora evaluation for multi-class wound classification. For this experiment, we considered the three classes of wounds that are common in the AZH and Medetec datasets, e.g., Diabetic, Pressure, and Venous wounds. This experiment is conducted on AZH, Medetec, Extended AZH, and Extended Medetec datasets.

We achieved this experiment in four phases, where at each phase, we trained the model on a training set of one dataset and tested it on a testing set of another dataset. For instance, in the first phase, we trained the model on the AZH dataset and then tested it on the Medetec dataset. Table 6 states the results of the cross-corpus experiment that demonstrate the generalization aptitude of our proposed SEEN-B4 model. The highest accuracy of 77% is achieved for this experiment when the model is trained on the Extended AZH dataset tested for the Extended Medetec dataset. It is worth noting that in the case of cross-corpora evaluation, the accuracies of the extended datasets are also higher compared to the original imbalance AZH and Medetec datasets. This indicates that the balanced dataset and more wound samples positively impact the cross-dataset evaluation. From Table 6, it is noticeable that the performance of our model drops for the cross-corpus experiment, despite having satisfactory performance on the individual datasets. The AZH and Medetec datasets are collected under different environments and conditions. For

example, the Medetec dataset includes stock images, while the AZH dataset images are gathered from the wound care center.

Table 6. Cross-Corpora Evaluation.

Training Dataset	Testing Dataset	Accuracy (%)	Balanced Accuracy (%)
AZH	Medetec	68.89	61.51
Medetec	AZH	59.15	59.68
Extended AZH	Extended Medetec	77.05	76.56
Extended Medetec	Extended AZH	58.06	58.06

Additionally, the images are captured from distinct devices and have different lighting conditions, color modes, background settings, varying angles, and distances from the capturing devices. All these variations among the images of the other datasets have made it challenging to achieve as good results, as has been attained for the individual datasets. Considering the cross-corpus scenario and the diversity among datasets, the results around 59–77% are encouraging, and supports the robustness of our method.

4.10. Limitations and Future Directions

In the domain of the wound classification, there are limited publicly available datasets for multi-class chronic wound classification. Additionally, the available datasets have a smaller number of samples. So, one of the future directions could be the development of public datasets that have more samples and wound classes compared to the existing datasets. Other future directions include the development of more robust generalized deep learning models that are capable of achieving improved performances on cross-corpora evaluation. Another important aspect is to develop explainable deep learning models that can better interpret the predictions of deep learning models for wound classification. In the future, we intend to further improve the classification performance of wound classification, along with enhancing the generalization power of the model. With this, the model will allow for a more accurate classification of wound type, which can further help caregivers to diagnose and to treat wounds based on a preliminary assessment.

5. Conclusions

This paper has proposed a novel and regularized fused Swish-ELU EfficientNet-B4 model to classify the different types of wound images accurately. The proposed model fuses the two customized EfficientNet-B4 models to extract better competent features, which are then classified via a regularized prediction module. Experimentation was performed on the Medetec dataset, the AZH dataset, and their extended versions, to demonstrate the efficacy of our SEEN-B4 model. To evaluate the generalization ability of the proposed model, we have also conducted a cross-corpora evaluation. We have generated and analyzed the heatmaps to strengthen the trustworthiness, and we demonstrate the explainability power of the SEEN-B4 framework. The results indicate that our proposed model can successfully classify the different wound types and outperform existing methods.

Author Contributions: Conceptualization, Z.A.A. and H.M.; methodology, Z.A.A., H.M. and R.M.; software, Z.A.A.; validation, Z.A.A., H.M. and R.M.; formal analysis, Z.A.A.; investigation, H.M.; resources, H.M. and R.M.; data curation, Z.A.A.; writing—original draft preparation, Z.A.A.; writing—review and editing, Z.A.A., H.M. and R.M.; visualization, Z.A.A.; supervision, H.M. and R.M.; project administration, H.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: We used the publicly available dataset, and links are already provided in the Reference section.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chronic Wounds. Available online: https://www.ncbi.nlm.nih.gov/books/NBK482198/ (accessed on 23 March 2023).
2. Chronic Wounds. Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9117969/ (accessed on 23 March 2023).
3. Chronic Wounds. Available online: https://www.healogics.com/wound-care-patient-information/category/common-chronic-wounds/ (accessed on 23 March 2023).
4. Chronic Wounds. Available online: https://www.fortunebusinessinsights.com/industry-reports/chronic-wound-care-market-100222 (accessed on 23 March 2023).
5. Zhang, R.; Tian, D.; Xu, D.; Qian, W.; Yao, Y. A survey of wound image analysis using deep learning: Classification, detection, and segmentation. *IEEE Access* **2022**, *10*, 79502–79515. [CrossRef]
6. Godeiro, V.; Neto, J.S.; Carvalho, B.; Santana, B.; Ferraz, J.; Gama, R. Chronic wound tissue classification using convolutional networks and color space reduction. In Proceedings of the 2018 IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP), Aalborg, Denmark, 17–20 September 2018; pp. 1–6. [CrossRef]
7. Blanco, G.; Traina, A.J.; Traina, C., Jr.; Azevedo-Marques, P.M.; Jorge, A.E.; de Oliveira, D.; Bedo, M.V. A superpixel-driven deep learning approach for the analysis of dermatological wounds. *Comput. Methods Programs Biomed.* **2020**, *183*, 105079. [CrossRef] [PubMed]
8. Wang, Y.; Ke, Z.; He, Z.; Chen, X.; Zhang, Y.; Xie, P.; Li, T.; Zhou, J.; Li, F.; Yang, C.; et al. Real-time burn depth assessment using artificial networks: A large-scale, multicentre study. *Burns* **2020**, *46*, 1829–1838. [CrossRef] [PubMed]
9. Gamage, C.; Wijesinghe, I.; Perera, I. Automatic Scoring of Diabetic Foot Ulcers through Deep CNN Based Feature Extraction with Low Rank Matrix Factorization. In Proceedings of the 2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE), Athens, Greece, 28–30 October 2019; pp. 352–356. [CrossRef]
10. Zhao, X.; Liu, Z.; Agu, E.; Wagh, A.; Jain, S.; Lindsay, C.; Tulu, B.; Strong, D.; Kan, J. Fine-grained diabetic wound depth and granulation tissue amount assessment using bilinear convolutional neural network. *IEEE Access* **2019**, *7*, 179151–179162. [CrossRef] [PubMed]
11. Al-Garaawi, N.; Ebsim, R.; Alharan, A.F.; Yap, M.H. Diabetic foot ulcer classification using mapped binary patterns and convolutional neural networks. *Comput. Biol. Med.* **2022**, *140*, 105055. [CrossRef] [PubMed]
12. Toofanee, M.S.A.; Dowlut, S.; Hamroun, M.; Tamine, K.; Petit, V.; Duong, A.K.; Sauveron, D. DFU-SIAM a Novel Diabetic Foot Ulcer Classification with Deep Learning. *IEEE Access* **2023**, *11*, 98315–98332. [CrossRef]
13. Khosa, I.; Raza, A.; Anjum, M.; Ahmad, W.; Shahab, S. Automatic Diabetic Foot Ulcer Recognition Using Multi-Level Thermographic Image Data. *Diagnostics* **2023**, *13*, 2637. [CrossRef] [PubMed]
14. Aldughayfiq, B.; Ashfaq, F.; Jhanjhi, N.Z.; Humayun, M. Yolo-based deep learning model for pressure ulcer detection and classification. *Healthcare* **2023**, *11*, 1222. [CrossRef] [PubMed]
15. Oliveira, B.; Torres, H.R.; Morais, P.; Baptista, A.; Fonseca, J.; Vilaça, J.L. Classification of chronic venous disorders using an ensemble optimization of convolutional neural networks. In Proceedings of the 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Glasgow, UK, 11–15 July 2022; IEEE: Piscataway, NJ, USA; pp. 516–519. [CrossRef]
16. Alzubaidi, L.; Fadhel, M.A.; Oleiwi, S.R.; Al-Shamma, O.; Zhang, J. DFU_QUTNet: Diabetic foot ulcer classification using novel deep convolutional neural network. *Multimed. Tools Appl.* **2020**, *79*, 15655–15677. [CrossRef]
17. Das, S.K.; Roy, P.; Mishra, A.K. DFU_SPNet: A stacked parallel convolution layers based CNN to improve Diabetic Foot Ulcer classification. *ICT Express* **2022**, *8*, 271–275. [CrossRef]
18. Abubakar, A.; Ugail, H.; Bukar, A.M. Assessment of human skin burns: A deep transfer learning approach. *J. Med. Biol. Eng.* **2020**, *40*, 321–333. [CrossRef]
19. Goyal, M.; Reeves, N.D.; Rajbhandari, S.; Ahmad, N.; Wang, C.; Yap, M.H. Recognition of ischaemia and infection in diabetic foot ulcers: Dataset and techniques. *Comput. Biol. Med.* **2020**, *117*, 103616. [CrossRef] [PubMed]
20. Hüsers, J.; Hafer, G.; Heggemann, J.; Wiemeyer, S.; Przysucha, M.; Dissemond, J.; Moelleken, M.; Erfurt-Berge, C.; Hübner, U.H. *Automatic Classification of Diabetic Foot Ulcer Images: A Transfer-Learning Approach to Detect Wound Maceration*; IOS Press: Amsterdam, The Netherlands, 2022. [CrossRef]
21. Malihi, L.; Hüsers, J.; Richter, M.L.; Moelleken, M.; Przysucha, M.; Busch, D.; Heggemann, J.; Hafer, G.; Wiemeyer, S.; Heidemann, G.; et al. Automatic Wound Type Classification with Convolutional Neural Networks. *Adv. Inform. Manag. Technol. Healthc.* **2022**, *295*, 281. [CrossRef]
22. Rostami, B.; Anisuzzaman, D.M.; 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. [CrossRef] [PubMed]
23. Sarp, S.; Kuzlu, M.; Wilson, E.; Cali, U.; Guler, O. The enlightening role of explainable artificial intelligence in chronic wound classification. *Electronics* **2021**, *10*, 1406. [CrossRef]

24. Anisuzzaman, D.M.; Patel, Y.; Rostami, B.; Niezgoda, J.; Gopalakrishnan, S.; Yu, Z. Multi-modal wound classification using wound image and location by deep neural network. *Sci. Rep.* **2022**, *12*, 20057. [CrossRef] [PubMed]
25. Huang, P.H.; Pan, Y.H.; Luo, Y.S.; Chen, Y.F.; Lo, Y.C.; Chen, T.P.C.; Perng, C.K. Development of a deep learning-based tool to assist wound classification. *J. Plast. Reconstr. Aesthetic Surg.* **2023**, *79*, 89–97. [CrossRef]
26. Anisuzzaman, D.M.; Patel, Y.; Niezgoda, J.; Gopalakrishnan, S.; Yu, Z. Wound Severity Classification using Deep Neural Network. *arXiv* **2022**, arXiv:2204.07942. [CrossRef]
27. Tan, M.; Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In Proceedings of the 36th International Conference on Machine Learning, Long Beach, CA, USA, 10–15 June 2019; pp. 6105–6114.
28. Clevert, D.A.; Unterthiner, T.; Hochreiter, S. Fast and accurate deep network learning by exponential linear units (elus). *arXiv* **2015**, arXiv:1511.07289. [CrossRef]
29. Selvaraju, R.R.; Das, A.; Vedantam, R.; Cogswell, M.; Parikh, D.; Batra, D. Grad-CAM: Why did you say that? *arXiv* **2016**, arXiv:1611.07450. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.