Intelligence-Based Medicine 8 (2023) 100128

Contents lists available at ScienceDirect

Intelligence-Based Medicine

journal homepage: www.sciencedirect.com/journal/intelligence-based-medicine

DFU_MultiNet: A deep neural network approach for detecting diabetic foot ulcers through multi-scale feature fusion using the DFU dataset

Shuvo Biswas [a], Rafid Mostafiz [b,*], Bikash Kumar Paul [a,c], Khandaker Mohammad Mohi Uddin [d], Md Masudur Rahman [e], F.N.U. Shariful [f]

[a] Department of Information and Communication Technology, Mawlana Bhashani Science and Technology University, Bangladesh
[b] Institute of Information Technology, Noakhali Science and Technology University, Bangladesh
[c] Department of Software Engineering, Daffodil International University, Bangladesh
[d] Department of Computer Science and Engineering, Dhaka International University, Bangladesh
[e] Department of Information and Communications Engineering, Noakhali Science and Technology University, Bangladesh
[f] School of Computing, University of North Florida, Jacksonville, FL, USA

ARTICLE INFO

Keywords:
Diabetic foot ulcer (DFU)
VGG19
NasNetMobile
DenseNet201
Feature fusion
Computer aided diagnosis (CAD)

ABSTRACT

Diabetic foot ulcer (DFU) is a common problem among people with diabetes that can result in amputation of the affected limb. Modern DFU treatment and diagnosis methods are expensive and time-consuming. Today, the development of the computer-aided diagnosis (CAD) method makes it possible for pathologists to diagnose DFU more swiftly and accurately. This has led to a rise in interest in deep learning (DL) approaches based on CAD. In this study, we introduce a novel framework called "DFU_MultiNet," which focuses on the transfer learning approach to classify healthy and ulcer skin images using publicly available repositories. The proposed framework is developed to offer an efficient and robust method for DFU classification that determines the distinction between healthy and ulcerated skin. The proposed approach extracts features from foot samples using three well-known pre-trained CNN models: VGG19, DenseNet201, and NasNetMobile. Finally, these extracted results are merged through a summing layer to create a powerful hybrid network. Through obtaining impressive accuracy (99.06 %), precision (100.00 %), recall (98.18 %), specificity (100.00 %), F1-score (99.08 %), and AUC (99.09 %) the proposed "DFU_MultiNet" framework holds great potential as a diagnostic tool in healthcare and clinical settings.

1. Introduction

Diabetic foot ulcers (DFUs) are an ultimate exposure of diabetes, identified by foot injuries. According to reports, the global population of diabetic individuals was 151 million in the year 2000, surged to over 422 million in 2014, and is now estimated to be around 537 million in 2021 [1]. In the past two decades, there has been a significant 10.5 % increase in the prevalence of diabetes among adults aged 18 years and older [2].

It is worth noting that around 80 % of these diabetic patients reside in developing countries, which often lack adequate healthcare facilities and resources, leading to a lower level of awareness about patient health conditions [3]. A significant proportion of diabetic patients (ranging from 15 % to 25 %) suffer from DFUs, which can progress to a severe stage, necessitating lower limb amputation, hospitalization, and even death if left untreated [4,5]. Infection with DFUs is a common cause of limb or foot amputation [6].

It reduces patients' survival rates and diminishes the power of human life, impacting their ability to earn a livelihood and participate in social activities [7]. Conditions such as gangrene and tissue death from disease can lead to the need for amputation, and the problem of DFUs is expected to rise in the future [8]. The foot is a vital part of human body, but sadly, approximately a million patients who have high blood sugar will lose this vital part every year. After analyzing several studies, we observed that every 20 s, a diabetic foot undergoes an operation.

Doctors need thorough information for a correct diagnosis and better treatment of DFU. Traditional diagnostic techniques require a lot of manual labor and are prone to mistakes. Computer-assisted diagnostic

* Corresponding author.
E-mail addresses: it21620@mbstu.ac.bd (S. Biswas), rafid.iit@nstu.edu.bd (R. Mostafiz), bikash.k.paul@ieee.org (B.K. Paul), jilanicsejnu@gmail.com (K.M. Mohi Uddin), masudur@nstu.edu.bd (M.M. Rahman), n01501509@unf.edu (F.N.U. Shariful).

https://doi.org/10.1016/j.ibmed.2023.100128
Received 23 June 2023; Received in revised form 2 October 2023; Accepted 22 November 2023
Available online 28 November 2023
2666-5212/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

(CAD) techniques improve performance with minimal expenses. Moreover, recent advancements in wearable health and mobile technologies used to treat diabetes related complications. They can enhance the patient's quality of life and extend remission by detecting and managing detrimental inflammation and foot pressure [9]. Sometimes sensors and devices help to detect various types of signals, such as physical, chemical, and biological signals. They can record and measure these signals in an expertise process and widely used in current medical systems. When new sensors and sensor-related technologies are schematically developed, non-medical areas will adjust them for utilization in their business sectors. The evolution of new medical sensor points to a wider application of these devices in the healthcare sector [10]. Medical imaging [11–13] plays a critical role in diagnosing and treating various medical conditions. The performance of machine learning (ML) and deep learning (DL) approaches in this field heavily relies on the use of advanced feature extraction and selection methods that can accurately capture important visual characteristics such as color, shape, and size. Previous studies utilizing ML and CNN approaches have achieved remarkable outcomes in accurately diagnosing DFUs. However, there is still a need for further research to ensure the efficacy of these techniques in real-world settings with various functions. Accurate diagnosis and proper management of DFUs play a crucial role in improving the patient's prognosis. DFU management [14] involves various procedures such as vascular resection, wound removal, and infection treatment. The DFU treatment regimen depends on the type and condition of the foot wound. The challenges associated with DFU care involve a series of educational challenges that require the evaluation and comparison of various classification, detection, and segmentation [15] techniques to determine the pre-trained techniques and their potential applications [16].

Transfer learning refers to the reuse of a pre-trained CNN model. This technique is highly popular in DL due to its ability to train deep CNN networks with a small dataset. This technique works better for problems with large datasets. However, most real-world problems consist of large datasets (i.e., the medical sector) that cannot be tackled by traditional transfer learning. To address this major issue, the multi-scale transfer learning (MTL) technique is applied to solve these problems.

MTL enhances transfer learning by incorporating multiple pre-trained CNN models within the neural network, whereas traditional transfer learning uses only one pre-trained CNN model. These models process samples at various scales and capture features at different levels of detail. Once features are extracted from each model, they are harmoniously combined to create a comprehensive representation.

MTL is particularly beneficial for data scientists to extract features from the multiple pre-trained CNN models at multiple scales. Both MTL and transfer learning can be employed in medical image classification, treatment planning, disease diagnosis, and anomaly identification. In the healthcare sector, these techniques can be applied to improve diagnostic tools and increase their precision and dependability. Processing medical images at multiple scales enhances the ability to capture the fine features of lesions effectively. MTL contributes to early detection in medical images, facilitating prompt interventions, and ultimately enhancing patient outcomes. Additionally, these models excel at rapidly analyzing vast quantities of medical images, enabling clinicians to make quicker and more accurate decisions.

In this study, we proposed a hybrid "DFU_MultiNet" framework-based automatic DFU categorization system that is capable of distinguishing between normal and ulcer foot skin from the DFU dataset. The DFU data were first preprocessed and partitioned into train-test sets. Then the training set is fed into the multi-scale transfer learning (MTL) model. Three popular pre-trained CNN models, namely DenseNet201, NasNetMobile, and VGG19, make up this MTL model, which is applied to extract features from samples of foot skin. After that, we integrated all the extracted results through a summing layer and fine-tuned those using two dropout layers, two dense layers, and two batch normalization (BN) layers. Finally, the final dense layer is employed for the DFU classification task.

The crucial points of this research work are described as follows.

1. The "DFU_MultiNet" framework propose a multi-scale feature fusion technique which outperform the existing models.
2. Proposed framework demonstrates a remarkable performance on classifying all the ulcer and healthy images with omittable miss classification rate.
3. "DFU_MultiNet" provides a segmentation-free feature extraction technique on a large DFU dataset and show the proficiency in acquiring the high results in DFU dataset. In this technique, no manual feature extraction methods are necessary, distinguishing it from conventional ML approaches. On the other hand, it directly captures pertinent features from the entire image, eliminating the need for a distinct segmentation process.

The other parts of this paper have been structured into several sections. In Section 2, we described the literature review. After that, the proposed "DFU_MultiNet" framework has been explained in detail, including data pre-processing and MTL model-building strategies in Section 3. The experimental setup and consequences are presented comprehensively in Section 4. Section 5 contains a comparative analysis of the outcomes and suggestions for future direction. The final findings of this study are summarized in Section 6.

2. Literature review

Numerous studies have been published on the classification and detection of DFU images to distinguish between healthy skin and ulcer skin. Most of the authors propose ML and image processing approaches that analyze different features like patterns, hue, and morphological textures. The proposed work's performance depends on the adopted models and training approaches. Various types of DFU-related works are briefly described in this section.

Kaselimi et al. [17] (**2022**) introduced a comprehensive review of the existing research on the use of artificial intelligence (AI) in monitoring DFUs, highlighting the advantages of these methods while also acknowledging the challenges in implementing them effectively for remote patient care. Their analysis focused on the imaging methods and optical sensors employed to detect DFUs. The study contemplates both the characteristics of sensors and the physiological aspects of the patients. The image data source recommended a number of monitoring tactics, which limits the use of AI algorithms [18].

Thotad et al. [19] (**2022**) proposed a DL approach named EfficientNet, for the early forecast and detection of diabetic foot ulcers (DFU). EfficientNet was employed for a sample set whose image size was 844 feet, consisting of diabetic ulcers and healthy skin. In this approach, they built up a robust network by adjusting three crucial properties (resolution, width, and depth) of the CNN model to classify diabetic and normal feet. Their method attained excellent results compared to modern algorithms like GoogleNet, VGGNet (VGG16 and VGG19), and AlexNet. It gave the utmost accuracy, recall, precision, and f1-score of 98.97 %, 98 %, 99 %, and 98 %, respectively.

In another method, Juan et al. [20] (**2022**) proposed a novel deep (CNN) classifier named DFU_VIRNet, for an automatic DFU skin classification task. Furthermore, their method focused on estimation maps to identify the probability of risk areas responsible for developing DFU. Two types of samples, namely, visible and invisible, were fed to the proposed scheme for training and testing purposes. The DFU_VIRNet provided the highest AUC score (0.99301) and ACC (0.97750), beating the recent outcomes.

Doulamis et al. [21] (**2021**) proposed a valuable non-invasive device that utilizes photonic-based technology for the treatment of DFUs in diabetic patients. This device employs hyperspectral and thermal imaging to evaluate the ulcer's condition and estimates the biomarkers deoxyhemoglobin and oxyhemoglobin using the imaging technique.

Additionally, this novel device was improved by incorporating signal processing methods utilizing DL for improving pixel accuracy and reducing noise using super-resolution approaches.

Das et al. [22] (**2021**) proposed a unique framework (DFU_SPNet), which was constructed from stacked parallel (SP) convolution layers. DFU_SPNet employed three distinct kernel size modules of SP convolution layers to extract the feature map. Obtaining ACC (97.4 %), the DFU_SPNet outperformed the existing ultra-modern works after evaluating the DFU test dataset utilizing the optimizer (SGD) with a learning rate (1-e2).

Alzubaidi et al. [23] (**2021**) proposed four hybrid CNN models for classifying abnormal skin vs. normal skin. The models they developed incorporated traditional CNN layers along with parallel convolutional layers (PCL). Each model has six modules of PCL, but the range of the PCL branches is from 2 to 5. All models extracted the features from the same input, utilizing PCL with different kernel sizes, and then merged the extracted results. Among the models tested, the model with four branches achieved the highest F1 score (95.8 %).

Alzubaidi et al. [24] (**2020**) proposed a study that utilized a 754-foot DFU dataset from both healthy and diabetic ulcer patients. To automate the categorization of these images, a deep CNN named DFU_QUTNet was proposed, which differed from conventional CNNs. DFU_QUTNet was wider but not necessarily deeper. This approach improved the gradient propagation problem since errors were remitted over several different channels.

Tan and Le et al. [25] (2019) demonstrated that optimizing network resolution, width, and depth can enhance performance and revealed an extensive description of model scaling. They introduced a new scaling mechanism that utilized a compound coefficient to uniformly scale these parameters. By applying this mechanism to a baseline model, they constructed the EfficientNets family of CNN models. This study surpassed other existing ConvNets in accuracy, efficacy, and speed of inference while obtaining an impressive ACC (84.3 %) on ImageNet [26].

Manu Goyal et al. [27] (2017) endorsed the potential of conventional computer vision (CCV) features as a practical and cost-effective method for detecting foot ulcers in diabetic patients. They developed a new CNN-based framework called DFUNet to differentiate between DFUs and healthy skin by extracting an image feature map. DFUNet achieved the AUC (0.962) by applying a 10-fold cross-validation approach. It demonstrated better performance than utilizing traditional ML and DL approaches.

Wang et al. [28] (2017) employed a special capture box to capture DFU images and, on the other hand, identified the exact location of DFU by employing a two-stage SVM classification approach. These two stages were: (i) the segmentation stage, which extracts super-pixels; and (ii) the feature extraction stage, which extracts important features from the image.

Manu et al. [29] (2017) implemented a method for segmenting DFUs from whole-foot images. Though the method exhibited promising results, the approach had limitations. The first limitation was the impracticality of using a box surface for capturing DFU images due to concerns about potential infection risks in a healthcare setting, which emphasized the importance of exploring alternative approaches for image collection and analysis. The second was its inability to effectively process a large DFU dataset. Table 1 describes an overview of the published paper.

Table 1
A summery table of all the approaches.

Paper	Dataset size	Approach	Performance
Thotad et al. [19]	1688	EfficientNet	Accuracy = 98.97 %, Precision = 99 %, Recall = 98 %, F1-score = 98 %
Juan et al. [20]	13,200	DFU_VIRNet	Accuracy = 0.9775, Sensitivity = 0.98167, Specificity = 0.97333, Precision = 0.97355, Recall = 0.98167, F1-score = 0.97759, AUC = 0.99301
Doulamis et al. [21]	_	Hyperspectral and thermal imaging techniques	Early prediction and prognosis of a DFU, understanding the effect of the biomarkers on DFU
Das et al. [22]	3827	DFU_SPNet	Accuracy = 0.964, Sensitivity = 0.984, Specificity = 0.951, Precision = 0.926, Recall = 0.984, F1-score = 0.954, AUC = 0.974
Alzubaidi et al. [23]	17,053	Hybrid CNN	Precision = 97.3 %, Recall = 94.5 %, F1-score = 95.8 %
Alzubaidi et al. [24]	17,053	DFU_QUTNet+ SVM	Precision = 0.954, Recall = 0.936, F1-score = 0.945
Tan and Le et al. [25]	ImageNet	ConvNet	Accuracy = 84.3 %
Manu Goyal et al. [27]	22,777	DFUNet	Accuracy = 0.925, Sensitivity = 0.934, Specificity = 0.911, Precision = 0.945, F1-score = 0.939, AUC = 0.961
Wang et al. [28]	100	SVM	Sensitivity = 73.3 %, Specificity = 94.6 %
Manu et al. [29]	705	FCN-16s	Dice for ulcer region (UR) = 0.794 and surrounding skin (SS) = 0.851, Specificity for SS = 0.994, Sensitivity for UR = 0.789 and SS = 0.874 and MCC for UR = 0.785 and SS = 0.852

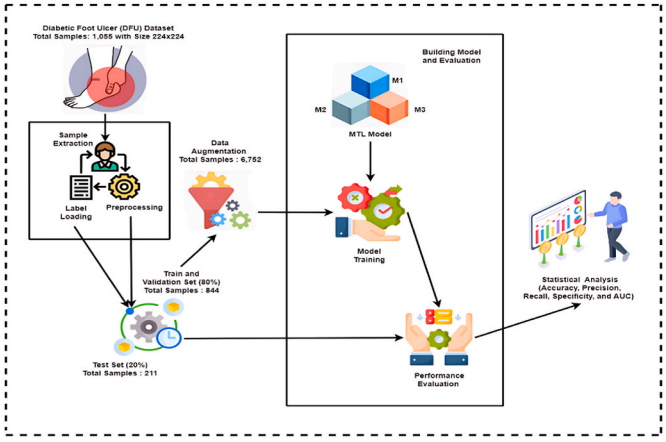


Fig. 1. Proposed DFU_MultiNet Framework for DFU classification tasks M1, M2, and M3 are three CNN models: DenseNet201, VGG19, and NasNetMobile, respectively. These three models built a hybrid MTL (multi-scale transfer learning) model.

3. Methodology

Developing a precise diagnosis system for classifying DFU is a challenging task for clinical purposes. This section explains how we proposed a hybrid "DFU_MultiNet" framework for the DFU classification challenge by merging various pre-trained CNN models. Fig. 1 **shows the framework of the proposed** "DFU_MultiNet". The framework performs the following operations: extracting DFU samples, loading labels, sample preprocessing, splitting the DFU dataset, data augmentation, training the MTL model, and finally, analyzing statistical parameters utilizing the test set. The overall classification task of our framework is shown step by step in Algorithm 1.

Algorithm 1 Automated diabetic foot ulcer (DFU) classification and detection.

Input: *DFU Training dataset γ1 (70%), Validation dataset γ2 (10%), and Testing dataset γ3 (20%)*

β = batch size

σ = epochs

λ = optimizer

η = learning rate

ε = the number of samples converted into mini-batch size

Output: *Ω = weight of pre-trained CNN algorithms*

Begin:

1: Convert each sample in the training dataset into a size of 224x224.

2: Apply the data augmentation method to enhance the sample size.

3: Extract the feature maps from the sample utilizing three pre-trained CNN algorithms, namely, VGG19, NasNetMobile, and DenseNet201.

4: Merge the extracted maps by applying the concatenation layer.

5: Initialize four fine-tuned CNN layers: batch normalization, dense, dropout, and softmax.

6: Set the training parameters: η, σ, λ, β, and ε.

7: Train the ''DFU_MultiNet'' framework and calculate the primary weights.

8: ***for*** *σ = 1 to σ* ***do***

9: Select a mini-batch size ε.

10: Forward propagation and evaluation of the binary loss function.

11: Backpropagation and improving the weight Ω.

12: ***end for***

End

3.1. DFU dataset

In this study, the working dataset named DFU dataset was collected from publicly accessible online repository [30], which contains four folders that are original images, patches, test set and transfer-learning images. To train and test our hybrid model, we selected the patch folder from these four folders, which included a total of 1055 skin patches. Out of these patches, 512 were identified as abnormal (ulcers), while the remaining 543 were classified as normal (healthy skin). Fig. 2 displays sample of skin patches. The DFU dataset is split into two sets, train (80 %) and test (20 %), with the help of the train_test_split function. The train_test_split function is imported from the "sklearn.model_selection" package in Python. Then, for the validation set, 10 % of the data is split from the train set with the help of the train_test_split function. Finally, the DFU dataset (i.e., 1055 samples) has been divided into three distinct phases: 70 % (i.e., 760 samples) for training, 20 % (i.e., 211 samples) for testing, and 10 % (i.e., 84 samples) for validation. Table 2 provides further details regarding the DFU dataset.

Table 2
Detail information about working data before applying augmentation techniques.

Dataset	Label	Training	Validation	Testing
DFU	Healthy skin	390	43	110
	Ulcer skin	370	41	101
	Total samples	760	84	211

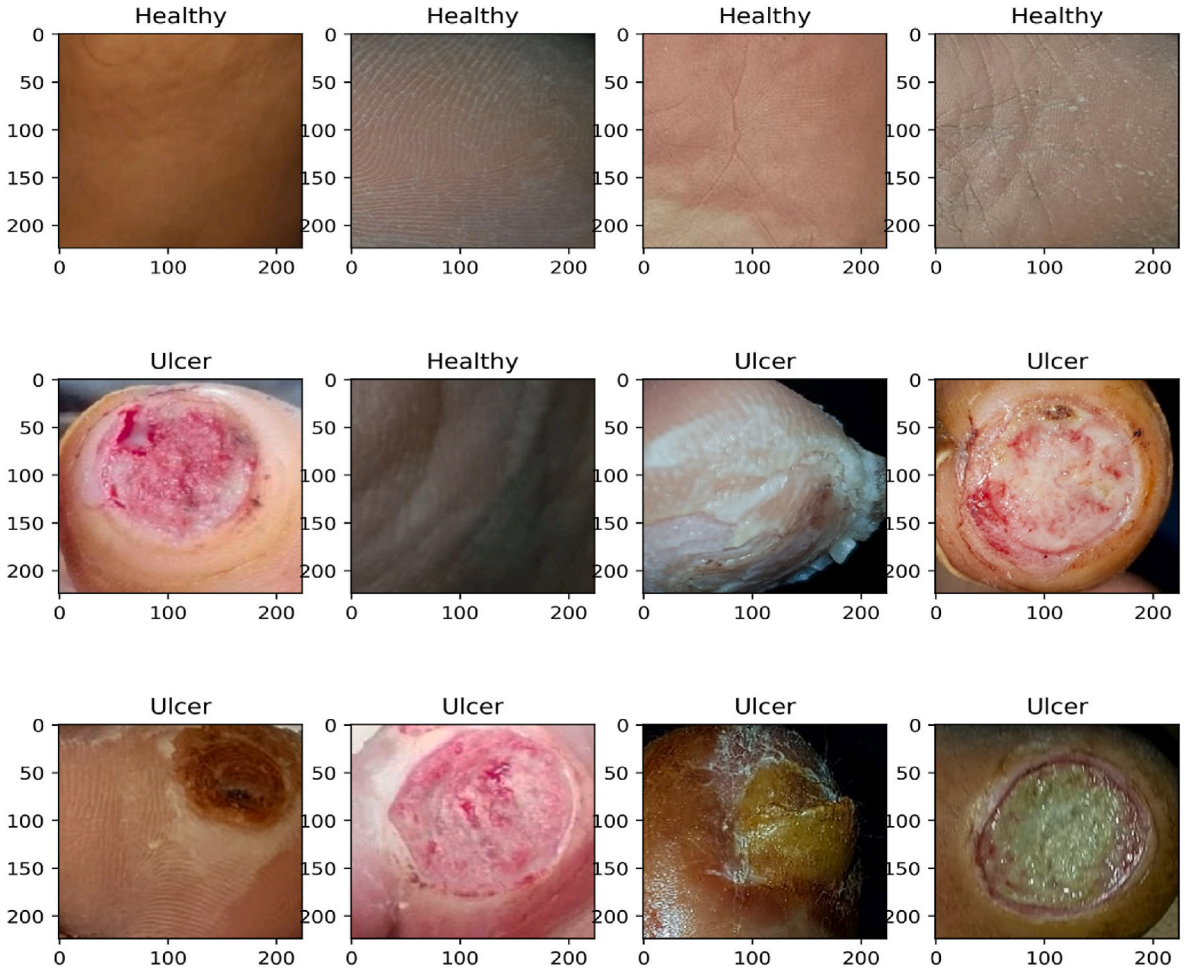


Fig. 2. Sample skin images (Healthy and Ulcer) of the DFU dataset.

3.2. Data preprocessing

Before inputting the DFU images into the multi-scale transfer learning (MTL) model, various pre-processing stages were implemented. According to transfer learning principles, each image in the DFU dataset was in.jpg format with a size of 224 × 224 pixels and channel RGB. The images were converted into Numpy arrays to enable faster training and reduce memory usage. Additionally, we have also shuffled the dataset for training unordered samples. The DFU dataset has been divided into distinct phases (see Fig. 3(a)), with 70 % (i.e., 760 samples) allocated to training, 20 % (i.e., 211 samples) to testing, and 10 % (i.e., 84 samples) to validation, respectively. The bar chart (Fig. 3(b)) indicates the number of samples after splitting the data into three phases. Deep networks require a large number of training samples due to their numerous parameters. This amount is increased by applying a powerful technique called data augmentation. This technique can serve various purposes, such as improving the performance of the DFU_MultiNet framework, handling overfitting matters, and enhancing the robustness of the model. Finally, applying this technique, we enhanced the DFU dataset from 1055 to 6963. In data augmentation techniques involving non-binary (true or false) parameters like rotation angle, shift, zooming, or shearing, these parameter values are usually selected randomly from predefined ranges or distributions. For instance, in image rotation, the rotation_range argument allows random selection of any degree between 0 and 360. When zooming an image, it is typically scaled within the range of [1 - zoom_range, 1 + zoom_range]. Shearing involves selecting a floating-point value from a uniform distribution in the range between 0 and 1. In contrast, flipping an image is binary, represented as True or False. All the augmentation parameters that were used in this study are given in Table 3 Table 4 reveals detailed information about the DFU dataset after the augmentation approach. Fig. 4 shows some examples of augmented images.

3.3. Building the MTL model

There is a growing interest in applying pre-trained neural networks to a wide range of tasks beyond their initial domains [31]. This is particularly demandable in the medical field, where obtaining sufficient labeled data for training DL networks can be challenging [32]. To address this, researchers have harnessed the power of networks trained on ImageNet [33], a vast database containing over 14 million (M) images in 20,000 categories but used 1.2 M images spanning more than 1000 classes for benchmarking. These classes may be abstract concepts, scenes, objects, and animals. The utilized pre-trained models in this experiment were pre-trained on this vast database and the weights of these networks are determined by this database.

In DL, these pre-trained networks have already acquired valuable representations and features from this vast database. These learned features enhance a model's proficiency in handling unseen data but this advantage is not obtained from the re-trained model. Conversely, training DL models from scratch necessitates computational resources and powerful hardware but these resources can be saved by using pre-trained models. These models are typically constructed upon well-structured architectures that have been greatly fine-tuned and tested. Given these advantages, in this study we opting pre-trained models rather than re-training them from scratch.

Table 3
Data augmentation techniques and parameters.

Number	Data techniques strategies	Parameter values
1	Zooming range	2
2	Rotation range	90
3	Shearing range	0.4
4	Width shift range	0.2
5	Height shift range	0.2
6	Horizontal flip	True
7	Vertical flip	True

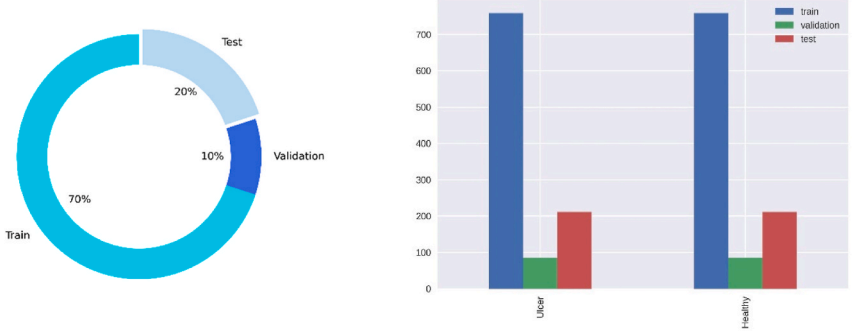
Table 4
Detail information about working data after applying augmentation techniques.

Dataset	Label	Training	Validation	Testing
DFU	Healthy skin	3120	344	110
	Ulcer skin	2960	328	101
	Total samples	6080	672	211

To build the MTL model, firstly three pre-trained models DenseNet201, NasNetMobile, and VGG19 are each fed the DFU color images with a size of 224 × 224 as input so that each model can extract the features of the images separately. Then each of the three GlobalAveragePooling2D layers is separately applied to each model, flattening the respective layers into a vector by averaging the features of each input. Subsequently, these individual vectors are combined into a unified vector by employing the concatenate layer. After that, the integrated features are fine-tuned using six CNN layers. Within these six layers, the first is the dropout layer (dropout rate 0.4), the second is the batch normalization layer, the third is the dense layer (128 units and ReLU activation function), the fourth is the dropout layer with a dropout rate of 0.6, the fifth is the batch normalization layer, and the final layer is the dense layer (2 units and softmax activation function). Fig. 5 depicts the MTL (multi-scale transfer learning) model for classifying DFU images. The MTL model contains 43,077,398 parameters after merging all of the collected features, which is roughly 2, 2, and 10 times more compared to the VGG19, DenseNet201, and NasNetMobile models. The basic explanation of these three adopted models and the fine-tuning approach are given in the following sub-sections.

3.3.1. DenseNet

Huang et al. (2017) [34], initially developed the best pre-trained



(a): Data splitting ratio. **(b):** Number of samples after splitting.

Fig. 3. (a) Data splitting ratio. Fig. 3 (b) Number of samples after splitting.

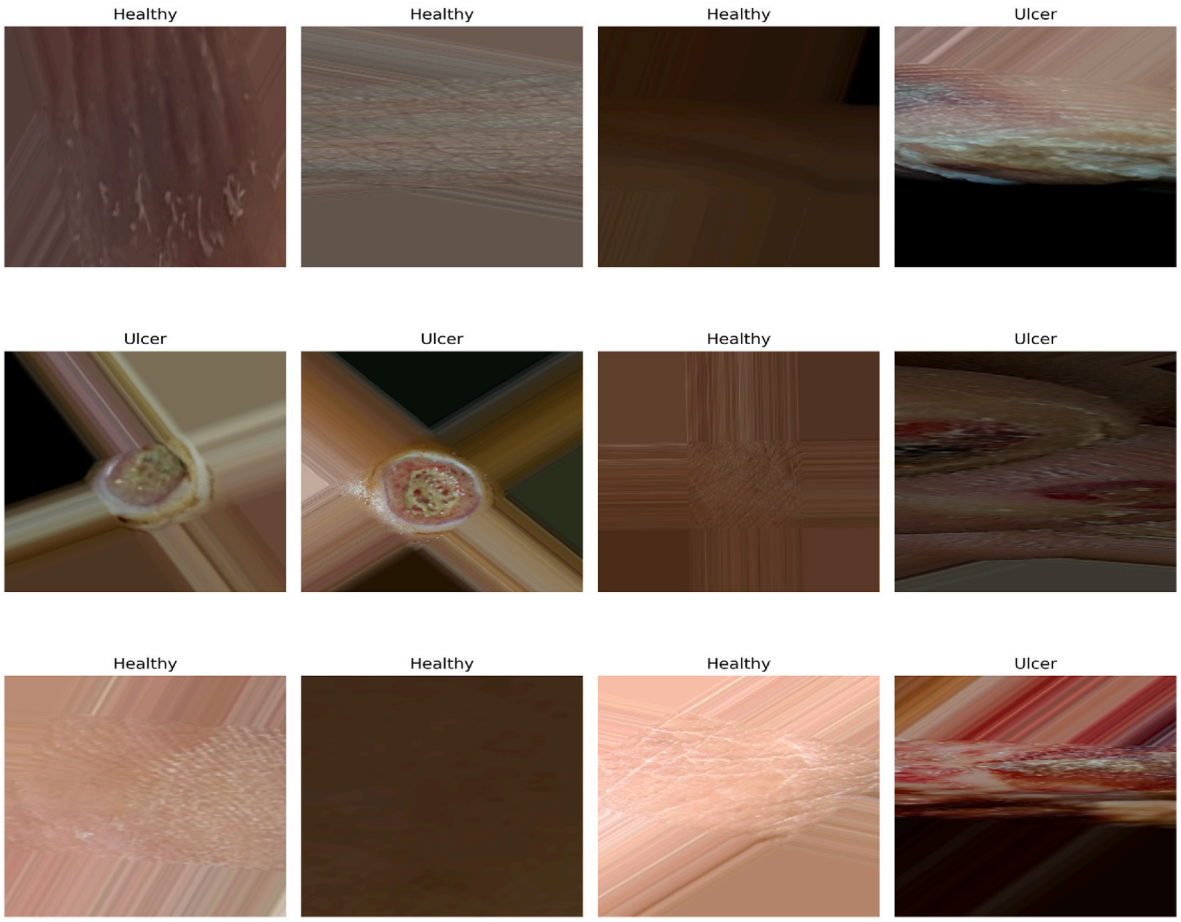


Fig. 4. Sample augmented images.

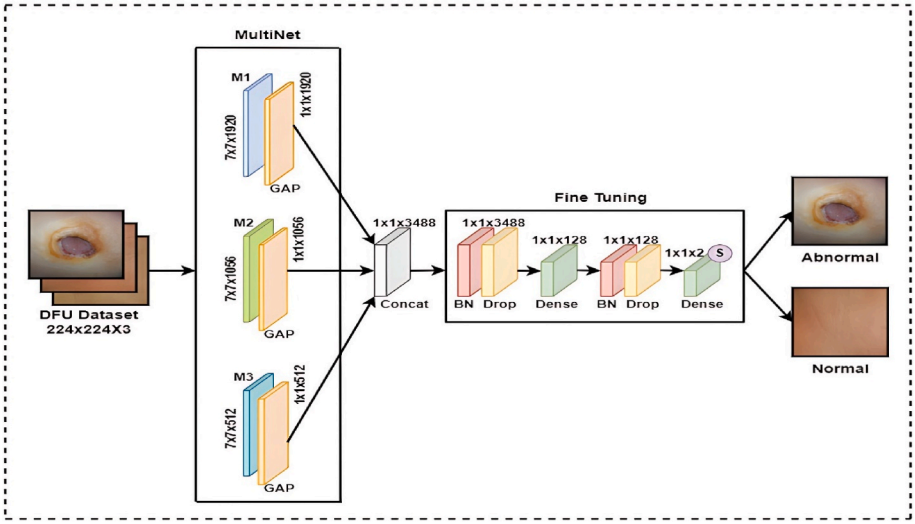


Fig. 5. MTL (multi-scale transfer learning) model to extract features from samples and merge them for diagnosis of ulcer skin. M1, M2, and M3 are three pre-trained DenseNet201, VGG19, and NasNetMobile models, respectively. GAP indicates the global average pooling layer, Concat indicates the concatenation layer, BN indicates the batch-normalization layer, and Drop indicates the dropout layer.

CNN classification model called DenseNet for obtaining the best accuracy on the ImageNet, CIFAR10, and CIFAR-100 datasets. This model was created using a feed-forwarding model similar to the ResNet model. Such a connection enables the architecture to exchange crucial data within the network, improving model performance and increasing the effectiveness of model training [35]. In this study, we employ DenseNet201 as our first feature detector network. This network has 201 deep neural layers in total, each of which is constructed to address overfitting difficulties while working with unordered samples. Additionally, after training this model, it contains a total of 18,321,984 parameters.

3.3.2. 3.3.2 NasNetMobile

Zoph et al. [36] (2018) first proposed NasNetMobile, our second feature extractor pre-trained model. On the CIFAR-10 dataset, the NasNetMobile model achieved a 2.4 % error rate using a novel regression method called ScheduledDropPath. In this study, this model is trained and tested on 224 × 224 DFU images using approximately 5.3 M training parameters. According to Saxena et al. [37] (2019), there exists an optimized model comprised of core building blocks that have been optimized through reinforcement learning. These blocks incorporate various pooling, convolution, and separable-convolution functions that enhance the overall reliability of the model.

3.3.3. 3.3.3 VGGNet

Simonyan et al. [38] (2014) proposed VGGNet, which achieved high performance in image localization and classification, ranking first and second, respectively, at the ILSVRC competition. Compared to the AlexNet architecture, VGGNet exhibited an impressive error rate (8.1 %) that is much better than the AlexNet architecture. For this study, we utilized VGG19 as the last feature extractor to build our DFU_MultiNet framework. The popular VGG19 model is organized by 16 convolution layers and three FC (fully connected) layers. The filter sizes for each convolution layer range from 64 to 512, with a 3 × 3 window size for all of them. Five blocks make up this model, with the first four convolution layers located in the first two blocks and the remaining twelve located in the next three. An MP (max pooling) layer after each block with a 2 × 2 window size detects the most important features from the modified activation maps [39]. The activation function (ReLU) is applied to each convolution layer. Finally, this model obtained a total of 20,024,384 parameters after training on the DFU dataset, which is more than the other two models.

3.4. Fine-tuning process

Fig. 5 illustrates how to integrate three pre-trained CNN models to build the MTL model for categorizing the DFU dataset, utilizing various FC (fully connected) layers. These three models employ GlobalAveragePooling2D simultaneously to flatten into a vector, which is done by computing the average value of the input images. A concatenate layer is then used to merge each vector into a single vector and fed through six additional layers with an activation function (softmax) for fine-tuning purposes for categorizing the DFU dataset. The explanations of each DL layers are given below.

The DL model faces a significant problem known as overfitting, which happens when it over trains on training data and performs poorly on test data [40]. We employ two dropout regularization layers to overcome the overfitting situation. During the DFU_MultiNet framework training, these layers excluded 40% and 20% of the samples, while also significantly improving training time. Additionally, such a procedure significantly speeds up the DFU dataset training task [41].

On the other hand, the inclusion of two BN (batch normalization) [42] layers is crucial for the success of our DFU_MultiNet framework. The main operations of this layer are to rescale and normalize the DFU samples, which makes the model more robust and reliable.

The dense layer also called the FC (fully connected) [43] layer connects all neurons between the two layers (previous and current). The main task of this special layer is to process input samples and generate the classification result. In our approach, we employ two FC layers, where the first use ReLU [44] and the second uses softmax as an activation function. This final layer predicts the length of the class and generates the DFU prediction. The softmax determines the most relevant features to predict the normal/ulcer class, whose outcome value ranges from 0 to 1, and triggers the neuron accordingly. It can be expressed as the following equation:

$$\mathrm{Softmax}(\mathrm{w})_{\mathrm{p}} = \frac{\exp(w_p)}{\sum_{m=1}^{n}\exp(w_p)} \quad (1)$$

The results of combining various pre-trained models with FC (fully connected) layers are reported in table (see Appendix A). This table was obtained during the construction of the "DFU_MultiNet" framework for DFU classification. That's why, in this framework, the final FC layer contains two neurons.

4. Dataset description, performance metrics, and results analysis

In the experimental setup, hyperparameters utilized in the study, and outcomes achieved by the "DFU_MultiNet" framework from the DFU dataset are presented in this section. Moreover, a comprehensive comparative discussion between the "DFU_MultiNet" framework and individual state-of-the-art CNN frameworks is conducted to evaluate the effectiveness of the proposed framework.

4.1. Dataset description

The DFU-dataset, which is available online [30], was used for training and testing. This repository comprises of four folders. The "original images" folder contains 493 images of different patients' feet with healthy feet and diabetic ulcers. The photos are from the diabetes center of the Nasiriyah Hospital in southern Iraq [24], and it's noteworthy that written consent and ethical approval were diligently obtained from all relevant patients and persons involved in the data collection process. These photos were taken by the experts with an iPad and a Samsung Galaxy Note 8 in a variety of lighting and viewing conditions. The "patches" folder contains 543 normal (healthy skin) and 512 abnormal (ulcer) skin patches, which were cropped from the samples of the "original images" folder with resolution 224 × 224 pixels. In this experiment, we used "patches" folder samples to train and test our model.

4.2. Experimental setup

The proposed "DFU_MultiNet" framework was developed with the help of Keras [45], for connecting Python [46] to the NN (neural network). The experimental setup with parameters described in Table 5.

4.3. Performance metrics

The performance evaluation of the "DFU_MultiNet" framework involved the use of various statistical parameters, such as accuracy (ACC), Kappa statistic, F1-score (FS), precision (PRE), Matthews correlation coefficient (MCC), specificity (SPE), sensitivity (SEN), and recall (REC). These parameters were calculated based on the values of false negative (FN), true positive (TP), true negative (TN), and false positive (FP) in the confusion matrix.

In the evaluation of the "DFU_MultiNet" framework, **TP** refers to the correct identification of positive foot skin. TN refers to the accurate identification of negative foot skin. On the other hand, FP denotes inaccurate identification of positive foot skin and FN refers to its incorrect identification of negative foot skin.

The followings are the performance metrics of the "DFU_MultiNet" framework:

Accuracy (ACC): ACC refers to the relationship between appropriately identified samples and the total samples of the DFU dataset.

Table 5
Experimental setup with parameters.

Item	Performance
Platform	Google Colab
GPU	Tesla K80
RAM	64 GB
CPU	Intel Core i5-12600K @ 3700 MHz

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

Recall (REC): REC refers to the ability of the "DFU_MultiNet" framework to successfully identify TP (true positive) samples by calculating the ratio of total positive samples in the DFU dataset.

$$Recall(REC) = \quad (3)$$

Specificity (SPE): SPE refers to the ability of the "DFU_MultiNet" framework to successfully identify TN (true negative) samples by calculating the division of total negative samples of the DFU dataset, also called the TNR (true negative rate).

$$Specificity(SPE) = \frac{TN}{TN + FP} \quad (4)$$

Precision (PRE): PRE refers to the ability of the "DFU_MultiNet" framework to successfully identify TP (true positive) samples by calculating the proportion of successfully predicted positive samples to all the predicted positive examples of the DFU dataset.

$$Precision(PRE) = \frac{TP}{TP + FP} \quad (5)$$

F1-Score (FS): The harmonic mean of REC and PRE is known as the FS.

$$F1Score(FS) = 2 * \frac{PRE * REC}{PRE + REC} \quad (6)$$

Matthews correlation coefficient (MCC): MCC is a statistical parameter that is applied for binary labeling. The value is bounded from −1 (worst outcome) to 1 (best outcome). It can be expressed as the following equation:

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (7)$$

Cohen's Kappa Coefficient (Kappa): Kappa is utilized to compare the predicted classes from the DFU_MultiNet framework with the actual classes in the DFU data. The value is bounded from −1 (worst outcome) to 1 (best outcome). It can be expressed as the following equation:

$$Kappa = \frac{Total\ ACC - Random\ ACC}{1 - Random\ ACC} \quad (8)$$

4.4. Training and parameter optimization

Fig. 6 presents the simulation outcomes of the proposed "DFU_MultiNet" framework, which are extracted during the training phase of the framework. To train the "DFU_MultiNet", specific hyper-parameter values were employed, as outlined in Table 6. Optimizer and gradient descent loss functions are two crucial parts for selecting hyperparameters during the framework training. In selecting the optimizer function for our framework, we opted for Adam [42] due to its ability to effectively manage sparse gradients in large datasets by combining the desirable properties of RMSProp and AdaGrad optimizers. As our model focus on binary classification, we considered a loss function named binary cross-entropy. Determining an appropriate learning rate is crucial to minimizing the loss function, which is very challenging. In the DL approach, a small learning rate makes the CNN classifier training slower; for this reason, the weight of the model updates is very minimal. To mitigate these issues, we set learning rate = 0.0001, epochs = 50, and batch size = 32 for demonstrating an effective model. Fig. 6 clearly shows that after training the 12th epoch, the "DFU_MeltiNet" framework offered to achieve 98.68 % training accuracy, 97.65 % validation accuracy, 6.17 % training loss, and 10.03 % validation loss, respectively. Fig. 6(a) further confirms that overfitting was not observed during the DFU_MeltiNet training process. Fig. 6(b) confirmed that the curve showed a rapid decrease in the loss value. However, some fluctuations occurred when selecting the narrow batch size.

Table 6
Training parameter with value for the "DFU_MultiNet" framework.

Number	Parameter	Value
1	Optimizer	adam
2	Learning Rate	0.0001
3	Minimum Learning Rate	1e-7
4	Decay	0.00001
5	Patience	5
6	Factor	0.2
7	Loss Function	binary_crossentropy
8	Metrics	accuracy
9	Batch Size	32
10	Epochs	50

Using a separate validation set for hyperparameter tuning is crucial as it allows for the selection of the best-performing model by systematically evaluating different hyperparameters. Additionally, it enables efficient searching for optimal hyperparameters and works as an indicator for detecting overfitting problems during the tuning process. If the test set is used for hyperparameter tuning, it may introduce data leakage, potentially affecting the model's performance. Furthermore, such tuning can result in overfitting to the test set. That's why hyperparameters shouldn't be tuned during the evaluation of the test set.

For hyperparameter tuning, the term "Factor" parameter signifies a scaling factor that is applied to adjust a hyperparameter's value. There isn't a fixed range for the "Factor" parameter. For instance, in DL, during tuning learning rates, the range for this parameter is [0.1, 10]. As we have integrated three models, we select smaller learning rates (0.0001) to induce finer updates to the weights during each iteration, which helps

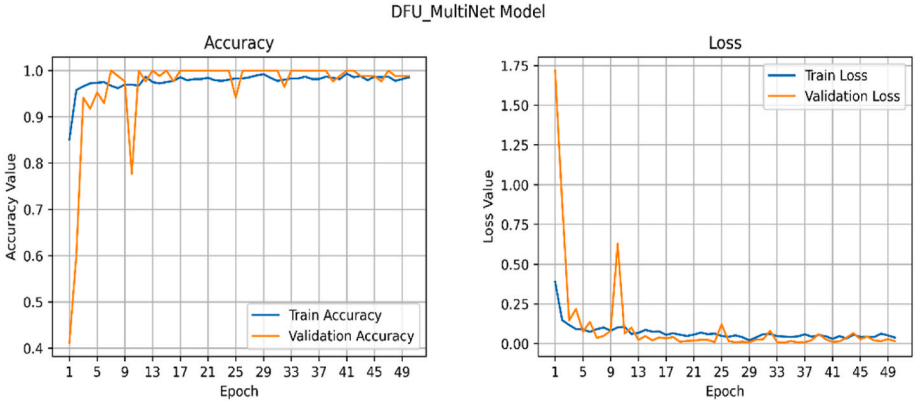


Fig. 6. Training improvement of the proposed DFU_MultiNet framework: (a) training versus validation accuracy plot (higher values indicate better performance), and (b) training versus validation loss plot (lower values indicate better performance).

to integrate models. When we train our model for too few epochs (like 10) then underfitting occurs, again while training for too many epochs (like 80) then overfitting occurs. For this reason, we select the epoch value 50 based on analyzing the validation accuracy and training loss (see Fig. 6). In DL, the batch size (BS) of 32 is a good initial point and a thumb rule [47]. It strikes a balance between computational efficiency and model accuracy. Large BSs can expedite training but may risk overfitting and reduced accuracy, whereas smaller BSs can be time-consuming and computationally expensive.

4.5. Results analysis

Fig. 7 presents the CM (confusion matrix) and ROC (receiver operating characteristic) curves for the DFU dataset, utilizing the "DFU_MultiNet" framework. The framework merged three renowned transfer learning algorithms, namely DenseNet201, VGG19, and NasNetMobile. By leveraging fusion features, the proposed approach effectively categorizes whether a diabetic foot is ulcerated or healthy. Fig. 7(a) reveals that the proposed "DFU_MultiNet" framework accurately classifies 101 ulcer skin images and 108 healthy skin images. Remarkably, the framework only misclassifies two instances of healthy skin. Notably, a significant advantage of the framework is its flawless performance in misclassifying no ulcer skin within the DFU dataset. The "DFU_MultiNet" framework demonstrates remarkable consistency and stability, as evidenced by the achieved AUC (0.99091) in Fig. 7(b). This high AUC score indicates the model's strong performance. Additionally, individual evaluations of all models on the DFU dataset further enhance the robustness of the "DFU_MultiNet" framework. Table 7 proves the superiority of the 'DFU_MultiNet' framework between the 'DFU_MultiNet' framework and five additional CNN models. The results indicate that the "DFU_MultiNet" framework attains impressive metrics such as precision of 1.00, recall of 0.982, f1-score of 0.991, kappa of 0.981, and MCC of 0.981, surpassing all other ultramodern models. Notably, DenseNet201 and VGG19 also demonstrate strong performance with accuracy values of 0.976 and 0.981, respectively. Meanwhile, VGG19 and MobileNet exhibit comparable performance in their ability to detect foot skin.

5. Discussion

The incidence of diabetic foot infections (DFI) and Diabetes-related problems can be caused by not obeying a healthy diet and no adequate safety precautions among the diabetics affected individuals. Ensuring proper guidance and caregiver to the diabetic patients may be imperative in addressing these challenges. Additionally, new techniques for diagnosis, therapy, and forecasting have been created as a result of applying technology to control diabetes. The proposed framework is based on the heterogeneous parallel ensemble DL architecture that has learned features in parallel from input samples through three pre-trained models (i.e. DenseNet201, VGG19, and NasNetMobile). This powerful technique can be leveraged across other clinical machine-learning applications. By utilizing this framework, healthcare professionals can make better decisions regarding patient enrollment in clinical trials, optimize drug development processes, and seamlessly integrate data from various sources. As a result, this network will play a vital role to enhance the efficiency and effectiveness of clinical research and healthcare systems. As shown in Table 7, the results from the suggested framework are more dependable and robust than those from the current models. This study suggests a novel method for classifying the skin of diabetic-affected feet by employing a hybrid "DFU_MultiNet" framework on DFU images. Table 8 provides a summary of the performance of the "DFU_MultiNet" framework in comparison to previous research that employed the same dataset but with various structures, depths, and parameters. Our approach, which combines predictions from multiple pre-trained models, offers several advantages for handling imbalanced datasets. It effectively mitigates overfitting and minimizes the possibility of noise impacting the minority label. This technique also assigns greater weight to minority class, resulting in improved classification for the imbalanced dataset. Since we used three different models, if one model fails to extract features from some data points in the dataset, others extract those features, which can enhance the ability to adapt to changes in the imbalanced dataset. Additionally, training these models in parallel on the same dataset significantly reduces the overall training time when dealing with an imbalanced dataset. Table 8 makes it obvious that, when compared to the earlier studies, the suggested framework offers the highest accuracy for the diagnosis of diabetic foot ulcers (DFU). Remarkably, the combination of all the pre-trained algorithms enables the framework to achieve a classification accuracy of 99 % for the DFU dataset. The ROC curve presented in Fig. 8 compares the performance of the DFU_MultiNet, a proposed framework, with several transfer learning networks such as DenseNet201, VGG16, VGG19, NasNetMobile, MobileNet, and DFU_MultiNet, using the same data partition. The results demonstrate that the DFU_MultiNet achieves an outstanding result over the standard transfer learning networks in accurately classifying ulcers versus healthy samples. Fig. 9 exhibits some of the diabetic-affected foot skin samples that are accurately predicted with the help of the novel DFU_MultiNet framework.

Adapting the "DFU_MultiNet" approach for clinical use involves several crucial steps and considerations to ensure its effectiveness and safety within a healthcare setting. Initially, the dataset must undergo annotation with ground truth labels to distinguish healthy feet from those with diabetic foot ulcers (DFUs). To enhance the transparency and understanding, various explainable AI techniques (e.g. LIME and SHAP) are employed to interpret the framework, providing clinicians with insights into the model's feature utilization for classification. Subsequently, integration of the "DFU_MultiNet" model into existing clinical systems or diagnostic tools utilized by healthcare professionals is imperative. The model's predictions must be presented in a clear and actionable format, ensuring that clinicians can readily comprehend and act upon the diagnostic results. Continuous monitoring of the model's performance is essential, encompassing diagnostic accuracy, treatment

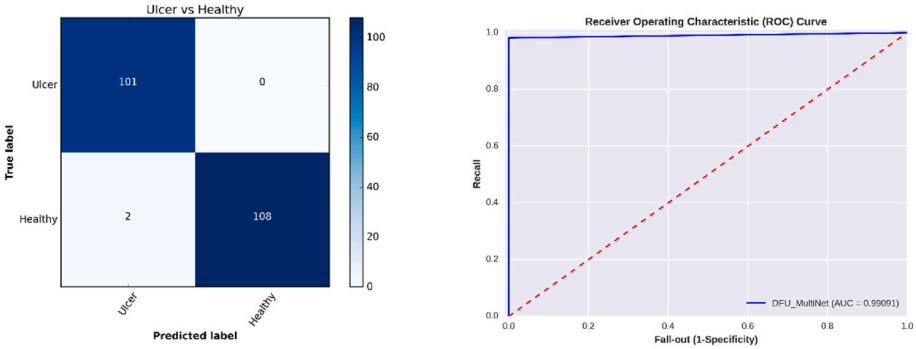


Fig. 7. (a) Confusion matrix (b) ROC curve for the "DFU_MultiNet" framework.

Table 7
Results attained from the "DFU_MultiNet" framework and five individual fine-tuned algorithms on DFU Dataset.

Model	Accuracy	Precision	Recall	$F1_{score}$	Specificity	AUC	Error Rate	MCC	Kappa
VGG19	0.867	**1.00**	0.745	0.854	**1.00**	0.873	0.133	0.764	0.737
VGG16	0.981	**1.00**	0.964	0.981	**1.00**	0.982	0.019	0.963	0.962
NasNetMobile	0.773	**1.00**	0.564	0.721	**1.00**	0.782	0.227	0.618	0.553
DenseNet201	0.976	**1.00**	0.955	0.977	**1.00**	0.977	0.024	0.954	0.953
MobileNet	0.867	0.894	0.845	0.869	0.891	0.868	0.133	0.736	0.735
DFU_MultiNet	**0.991**	**1.00**	**0.982**	**0.991**	**1.00**	**0.991**	**0.009**	**0.981**	**0.981**

Table 8
Comparison of "DFU_MultiNet" framework with existing techniques for DFU dataset.

Paper	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Classification Method	Training and validation Sample	Testing Sample
Thotad et al. (2022) [18]	98.97	99	98	98	EfficientNet	1350	338
Juan et al. (2021) [19]	97.8	97.4	98.2	97.8	DFU_VIRnet	12,600	600
K. Das et al. (2021) [21]	96.4	92.6	**98.4**	95.4	DFU_SPNet	3491	336
Alzubaidi et al. (2021) [22]	–	97.3	94.5	95.8	Hybrid CNN	16,731	322
Alzubaidi et al. (2020) [23]	–	95.4	93.6	94.5	DFU_QUTNet+ SVM	16,731	322
Goyal et al. (2017) [26]	92.5	94.5	–	93.9	DFUNet	22,605	172
Proposed Framework	**99.1**	**100**	98.2	**99.1**	**DFU_MultiNet**	6752	211

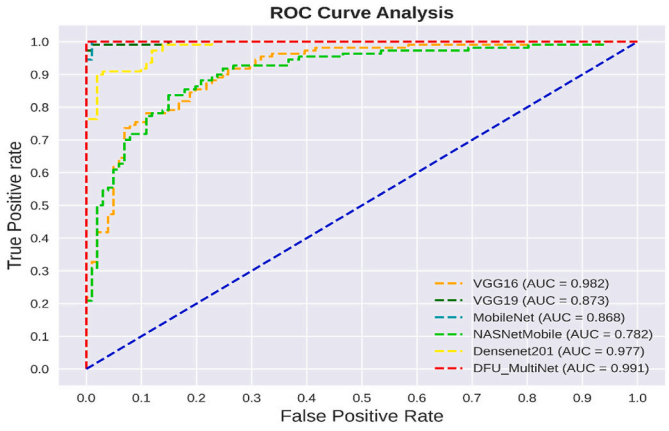


Fig. 8. ROC curve for several transfer learning algorithms.

planning, and patient outcomes assessment within real clinical scenarios.

In this clinical context, clinicians can employ the framework as a prospective Computer-Aided Diagnostic (CAD) tool. They can effortlessly input DFU images into the CAD tool's interface, where the "DFU_MultiNet" framework undertakes feature extraction by analyzing the images. These extracted features are then amalgamated through a summing layer, generating a comprehensive representation of the image for diagnostic evaluation and decision-making.

5.1. Limitations of the study

In this framework, it is feasible to diagnose whether the conferred sample is an ulcer or healthy only. It cannot provide a real-time assessment of pain severity or complexity levels.

To maximize the potential of this framework in clinical areas with limited computing power, future works should prioritize the development of distributed training algorithms and federated learning, alongside the creation of interpretable AI systems. Additionally, the development of DL models that operate efficiently on edge devices and mobile platforms will be crucial for clinical areas where there is not significant computing power available.

To test the performance of the DFU_MultiNet approach for DFU classification, we have trained this approach as well as five individual pre-trained models (i.e. DenseNet201, VGG19, VGG16, MobileNet, and NasNetMobile) on the same dataset with the same splitting ratio. These five pre-trained models are trained as follows: firstly, input images with a 224 × 224 size are fed into these models to produce a set of output feature maps by extracting features from the input image. A GlobalAveragePooling2D layer is then applied after each model to reduce the output feature map to a one-dimensional vector. This is done by taking the mean of all the values in the feature map. After that, six fine-tuning layers are added following the GlobalAveragePooling2D layer, one after another which follows in dropout - batch_normalization - dense and again dropout -batch_normalization - dense manner. These six layers are described in Section 3.3, which is also used to build the MTL model.

Though the proposed model exhibits good performance on this DFU dataset, it will provide a better and more accurate result on a larger dataset. During training on a large dataset, this framework will learn unique feature maps from different samples, which makes it a powerful ulcer detector tool. In healthcare and medical imaging, this powerful tool promises to detect disease more precisely and earlier, enabling better treatment planning and patient outcomes. In the future, this framework will be applicable for accurate quantification of different parameters within DFU samples, such as ulcer size, tissue density, volume, or growth rate. This quantitative data will facilitate ongoing patient monitoring and enhance treatment planning.

6. Conclusion

Regular foot examinations are essential for individuals with diabetes to identify any potential lesions, in addition to undergoing comprehensive assessments for peripheral arterial and neuropathy problems, as these conditions have the potential to result in the formation of ulcers or wounds. DFU can be prevented with regular foot exams, glucose control, patient concealment, suitable footwear, and timely treatment for pre-ulcerative infections. In this task, we have offered an innovative hybrid framework named "DFU_MultiNet" to diagnose diabetic-affected skin from foot samples more accurately and consistently. The "DFU_MultiNet" framework is based on a feature extraction and fine-tuning approach that permits various pre-trained CNN algorithms to extract and merge feature maps in parallel for DFU classification. It is developed in a balanced way that can control diverse DFU datasets. The outcomes of the experiments exhibit that the "DFU_MultiNet" framework, which surpasses both separate CNN pre-trained algorithms and all other modern algorithms reported in the published work, attains 99 %

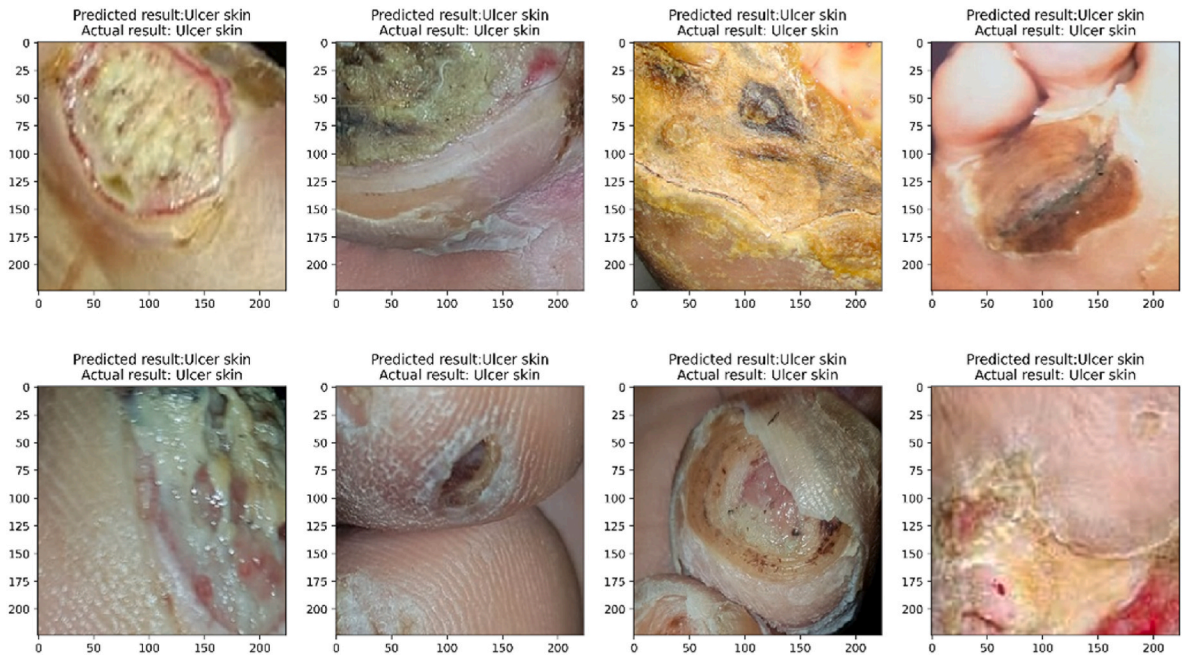


Fig. 9. "DFU_MultiNet" successfully tested on some DFU samples.

accuracy. This hybrid framework outperformed the most recent DFU classification methods. Considering the promising outcomes, we have strong confidence in the potential of our "DFU_MultiNet" framework as an excellent tool to aid doctors in efficiently detecting and diagnosing DFU. Additionally, it performs admirably in locating ulcer skins, improving the likelihood of survival.

In future studies, this hybrid framework should be expanded to detect and classify the DFU into ischemia, neuropathy, osteomyelitis, or Charcot arthropathy.

Data availability

The data supporting this study's findings are available from the corresponding author upon reasonable request.

Funding

None.

Ethical approval

Not required.

Consent to participate

Not required.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A

Details of the "DFU_MultiNet" framework.

Layer (type)	Output Shape	Param #	Connected to
input_1	(224, 224, 3)	0	
densenet201	(7, 7, 1920)	18321984	input_1[0][0]
NASNet	(7, 7, 1056)	4269716	input_1[0][0]
vgg19	(7, 7, 512)	20024384	input_1[0][0]
global_average_pooling2d	(1920)	0	densenet201[0][0]
global_average_pooling2d	(1056)	0	NASNet[0][0]
global_average_pooling2d	(512)	0	vgg19[0][0]
concatenate_4	(3488)	0	global_average_pooling2d[0][0] global_average_pooling2d_1[0][0] global_average_pooling2d_2[0][0]
dropout	(3488)	0	concatenate_4[0][0]
batch_normalization	(3488)	13952	dropout[0][0]
dense	(128)	446592	batch_normalization[0][0]
dropout_1	(None, 128)	0	dense[0][0]
batch_normalization_1	(None, 128)	512	dropout_1[0][0]

(*continued on next page*)

(*continued*)

Layer (type)	Output Shape	Param #	Connected to
dense_1	(None, 2)	258	batch_normalization_1[0][0]
Total params: 43,077,398			
Trainable params: 42,804,372			
Non-trainable params: 273,026			

References

[1] Diabetes [Online]. Available: https://www.who.int/health-topics/diabetes. [Accessed 29 September 2023].
[2] Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes. Diabetes Care May 2004;27(5):1047–53. https://doi.org/10.2337/diacare.27.5.1047.
[3] Aguirre F, Brown A, Cho NH, Dahlquist G, Dodd S, Dunning T, Hirst M, Hwang C, Magliano D, Patterson C, Scott C. IDF diabetes atlas. 2013.
[4] Mariam TG, Alemayehu A, Tesfaye E, Mequannt W, Temesgen K, Yetwale F, Limenih MA. Prevalence of diabetic foot ulcer and associated factors among adult diabetic patients who attend the diabetic follow-up clinic at the University of Gondar Referral Hospital, North West Ethiopia, 2016: institutional-based cross-sectional study. J Diabetes Res 2017. 2017.
[5] Almobarak AO, Awadalla H, Osman M, Ahmed MH. Prevalence of diabetic foot ulceration and associated risk factors: an old and still major public health problem in Khartoum, Sudan? Ann Transl Med 2017;5(17).
[6] Singh G, Gupta S, Chanda A. Biomechanical modelling of diabetic foot ulcers: a computational study. J Biomech 2021;127:110699.
[7] Pourkazemi A, Ghanbari A, Khojamli M, Balo H, Hemmati H, Jafaryparvar Z, Motamed B. Diabetic foot care: knowledge and practice. BMC Endocr Disord 2020; 20:1–8.
[8] Ghosh P, Valia R. Burden of diabetic foot ulcers in India: evidence landscape from published literature. Value Health 2017;20(9):A485.
[9] Najafi B, Reeves ND, Armstrong DG. Leveraging smart technologies to improve the management of diabetic foot ulcers and extend ulcer-free days in remission. Diabetes/metabolism research and reviews 2020;36:e3239.
[10] Formica D, Schena E. Smart sensors for healthcare and medical applications. Sensors 2021;21(2):543.
[11] Veredas FJ, Luque-Baena RM, Martín-Santos FJ, Morilla-Herrera JC, Morente L. Wound image evaluation with machine learning. Neurocomputing 2015;164: 112–22.
[12] Alzubaidi L, Fadhel MA, Al-Shamma O, Zhang J, Santamaría J, Duan Y. Robust application of new deep learning tools: an experimental study in medical imaging. Multimedia Tools and Applications; 2022. p. 1–29.
[13] Mostafiz R, Rahman MM, Uddin MS. Gastrointestinal polyp classification through empirical mode decomposition and neural features. SN Appl Sci Jun. 2020;2(6): 1143. https://doi.org/10.1007/s42452-020-2944-4.
[14] Vas PR, Kavarthapu V. Management of diabetic foot disease. In: Diabetic neuropathy. Elsevier; 2022. p. 235–58.
[15] Cassidy B, Kendrick C, Reeves ND, Pappachan JM, O'Shea C, Armstrong DG, Yap MH. Diabetic foot ulcer grand challenge 2021: evaluation and summary. In: Diabetic foot ulcers grand challenge: second challenge, DFUC 2021, held in conjunction with MICCAI 2021, strasbourg, France, september 27, 2021, proceedings. Cham: Springer International Publishing; 2022. p. 90–105.
[16] Yap MH, Hachiuma R, Alavi A, Brüngel R, Cassidy B, Goyal M, Zhu H, Rückert J, Olshansky M, Huang X, Saito H. Deep learning in diabetic foot ulcers detection: a comprehensive evaluation. Comput Biol Med 2021;135:104596.
[17] Najafi B, Mohseni H, Grewal GS, Talal TK, Menzies RA, Armstrong DG. An optical-fiber-based smart textile (smart socks) to manage biomechanical risk factors associated with diabetic foot amputation. J Diabetes Sci Technol 2017;11(4): 668–77.
[18] Kaselimi M, Protopapadakis E, Doulamis A, Doulamis N. A review of non-invasive sensors and artificial intelligence models for diabetic foot monitoring. Front Physiol 2022:2230.
[19] Thotad PN, Bharamagoudar GR, Kallur SS. Boosting-based machine learning approaches for diabetes prediction using Indian demographic and health survey-2021 data. 2023.
[20] Reyes-Luévano J, Guerrero-Viramontes JA, Romo-Andrade JR, Funes-Gallanzi M. DFU_VIRnet: a novel visible-infrared CNN to improve diabetic foot ulcer classification and early detection of ulcer risk zones. 2022.
[21] Doulamis A, Doulamis N, Angeli A, Lazaris A, Luthman S, Jayapala M, Silbernagel G, Napp A, Lazarou I, Karalis A, Hoveling R. A non-invasive photonics-based device for monitoring of diabetic foot ulcers: architectural/sensorial components & technical specifications. Inventions 2021;6(2):27.
[22] Das SK, Roy P, Mishra AK. DFU_SPNet: a stacked parallel convolution layers based CNN to improve Diabetic Foot Ulcer classification. ICT Express 2022;8(2):271–5.
[23] Alzubaidi LAITH, Abbood AA, Fadhel MA, Al-Shamma OMRAN, Zhang JIN. Comparison of hybrid convolutional neural networks models for diabetic foot ulcer classification. J Eng Sci Technol 2021;16(3):2001–17.
[24] Alzubaidi L, Fadhel MA, Oleiwi SR, Al-Shamma O, Zhang J. DFU_QUTNet: diabetic foot ulcer classification using novel deep convolutional neural network. Multimed Tool Appl 2020;79(21–22):15655–77.
[25] Liu Z, Mao H, Wu CY, Feichtenhofer C, Darrell T, Xie S. A convnet for the 2020s (2022. 2022. *URL:*, https://arxiv.org/abs/2201.03545.
[26] Tan M, Le Q. Efficientnet: rethinking model scaling for convolutional neural networks. In: International conference on machine learning. PMLR; 2019, May. p. 6105–14.
[27] Goyal M, Reeves ND, Davison AK, Rajbhandari S, Spragg J, Yap MH. Dfunet: convolutional neural networks for diabetic foot ulcer classification. IEEE Transactions on Emerging Topics in Computational Intelligence 2018;4(5):728–39.
[28] Wang L, Pedersen PC, Agu E, Strong DM, Tulu B. Area determination of diabetic foot ulcer images using a cascaded two-stage SVM-based classification. IEEE (Inst Electr Electron Eng) Trans Biomed Eng 2016;64(9):2098–109.
[29] Goyal M, Yap MH, Reeves ND, Rajbhandari S, Spragg J. Fully convolutional networks for diabetic foot ulcer segmentation. In: 2017 IEEE international conference on systems, man, and cybernetics (SMC). IEEE; 2017, October. p. 618–23.
[30] Dataset: diabetic foot ulcer (DFU). Available link: https://www.kaggle.com/laithjj/diabetic-foot-ulcer-dfu.
[31] Pan SJ, Yang Q. A survey on transfer learning. IEEE Trans Knowl Data Eng Oct 2010;22(10):1345–59.
[32] Shin HC, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers RM. Deep convolutional neural networks for computer-aided detection: cnn architectures, dataset characteristics and transfer learning. IEEE Trans Med Imag May 2016;35(5):1285–98.
[33] Russakovsky O, et al. ImageNet large scale visual recognition challenge. arXiv; Jan. 29, 2015 [Online]. Available, http://arxiv.org/abs/1409.0575. [Accessed 23 September 2023].
[34] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017. p. 4700–8.
[35] Hao W, Zhang Z. Spatiotemporal distilled dense-connectivity network for video action recognition. Pattern Recogn 2019;92:13–24.
[36] Zoph B, Vasudevan V, Shlens J, Le QV. Learning transferable architectures for scalable image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2018. p. 8697–710.
[37] Saxen F, Werner P, Handrich S, Othman E, Dinges L, Al-Hamadi A. Face attribute detection with mobilenetv2 and nasnet-mobile. In: 2019 11th international symposium on image and signal processing and analysis (ISPA). IEEE; 2019, September. p. 176–80.
[38] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 2014. *arXiv preprint arXiv:1409.1556*.
[39] Scherer D, Müller A, Behnke S. Evaluation of pooling operations in convolutional architectures for object recognition. In: Artificial neural networks–ICANN 2010: 20th international conference, thessaloniki, Greece, september 15-18, 2010, proceedings, Part III 20. Springer Berlin Heidelberg; 2010. p. 92–101.
[40] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res 2014;15 (1):1929–58.
[41] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. In: International conference on machine learning. pmlr; 2015, June. p. 448–56.
[42] Koushik J. Understanding convolutional neural networks. 2016. *arXiv preprint arXiv:1605.09081*.
[43] Dahl GE, Sainath TN, Hinton GE. Improving deep neural networks for LVCSR using rectified linear units and dropout. In: 2013 IEEE international conference on acoustics, speech and signal processing. IEEE; 2013, May. p. 8609–13.
[44] Chollet F. "Keras. Deep learning library for theano and tensorflow 2015;7(8):T1. URL: https://keras.io/k.
[45] Ketkar N. Deep learning with Python. Berkeley, CA: Apress; 2017. https://doi.org/10.1007/978-1-4842-2766-4.
[46] Kingma Diederik P, Ba Jimmy. Adam: A Method for Stochastic Optimization. 2014. arXiv preprint arXiv:1412.6980.
[47] Mostafiz R, Uddin MS, Uddin KMM, Rahman MM. COVID-19 along with other chest infection diagnoses using faster R-CNN and generative adversarial network. ACM Trans. Spatial Algorithms Syst. Sep. 2022;8(3):1–21. https://doi.org/10.1145/3520125.