Healthcare Analytics 3 (2023) 100143

Contents lists available at ScienceDirect

Healthcare Analytics

journal homepage: www.elsevier.com/locate/health

An in-depth analysis of Convolutional Neural Network architectures with transfer learning for skin disease diagnosis

Rifat Sadik [a], Anup Majumder [a], Al Amin Biswas [b,*], Bulbul Ahammad [a], Md. Mahfujur Rahman [b]

[a] Department of Computer Science and Engineering, Jahangirnagar University, Savar, Dhaka, Bangladesh
[b] Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh

ARTICLE INFO

Keywords:
Convolutional Neural Network (CNN)
Skin disease
Deep learning
Xception
MobileNet
Transfer Learning (TL)

ABSTRACT

Low contrasts and visual similarity between different skin conditions make skin disease recognition a challenging task. Current techniques to detect and diagnose skin disease accurately require high-level professional expertise. Artificial intelligence paves the way for developing computer vision-based applications in medical imaging, like recognizing dermatological conditions. This research proposed an efficient solution for skin disease recognition by implementing Convolutional Neural Network (CNN) architectures. Computer vision-based applications using CNN architectures, MobileNet and Xception, are used to construct an expert system that can accurately and efficiently recognize different classes of skin diseases accurately and efficiently. The proposed CNN architectures used a transfer learning method in which models are pre-trained on the Imagenet dataset to discover more features. We also evaluated the performance of our proposed approach with some of the most popular CNN architectures: ResNet50, InceptionV3, Inception-ResNet, and DenseNet, thus establishing a comparison to set up a benchmark that will ratify the essence of transfer learning and augmentation. This study uses data from two separate data sources to collect five different types of skin disorders. Different performance evaluation indicators, including accuracy, precision, recall, and F1-score, are calculated to verify the success of our technique. The experimental results revealed the effectiveness of our proposed approach, where MobileNet achieved a classification accuracy of 96.00%, and the Xception model reached 97.00% classification accuracy with transfer learning and augmentation. Moreover, we proposed and implemented a web-based architecture for the real-time recognition of diseases.

1. Introduction

1.1. Background

Skin is the most vital and sensitive organ in the human body, shielding against heat, injury, and infections. Unfortunately, the skin condition is sometimes disrupted due to bacterial and viral infection, fungus, lack of a strong immune system, and genetic imbalances. In many cases, diseases caused by those factors have macabre effects on human life. In addition, some skin diseases are contagious, risking not only individuals but also others related to the infected. Statistics [1] reported that over 100 million people all over the world are suffering from different types of skin indispositions; the most frequent skin disorders are Atopic dermatitis, Eczema, Herpes, Nevus, Warts, Ringworm, Chickenpox, and Melanoma, etc. American Cancer Society reported [2] that, by the end of the year 2020, 100,350 new melanoma cases will be reported and diagnosed, and almost 6850 people are about to die because of melanoma.

Most skin diseases have revealing symptoms such as rash, ulcers, lesions, moles, etc. However, the diagnosis of skin diseases faces some difficulties. The most common obstacle is that many skin conditions have similarities between them that are not distinguishable visually. Besides, symptoms are constantly changing over a long process. Even physicians are bound to visual imperfections due to the lighting conditions of the environment, the skin color of the patient, and their professional experience. In most cases, early detection of skin diseases reduces the risk factors. The mortality rate of some diseases with a high mortality rate can be reduced to 90% if diagnosed in the early stage [3].

1.2. Motivation

Researchers are actively investigating methods to develop skin disease recognition systems. Many studies have utilized image processing techniques incorporating statistical analysis to extract information

* Corresponding author.
E-mail addresses: rifat.sadik.rs@gmail.com (R. Sadik), anupmajumder@juniv.edu (A. Majumder), alaminbiswas.cse@gmail.com (A.A. Biswas), bulbul@juniv.edu (B. Ahammad), mrrajuiit@gmail.com (Md.M. Rahman).

https://doi.org/10.1016/j.health.2023.100143
Received 11 June 2022; Received in revised form 25 November 2022; Accepted 24 January 2023

2772-4425/© 2023 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

about skin conditions [4–8]. Researchers were trying to recognize skin diseases by analyzing textures, structures, and colors in these approaches. Methods like Self-organizing Map (SOM), Radial Basis Function (RBF), Gray Level Co-occurrence Matrix (GLCM), etc., were used for such approaches. But all these methods lack in terms of precision and accuracy since these methods require sufficient data, good coverage of the input space, and high dependence on texture features such as contrast, correlation, entropy, etc.

In recent times, Artificial Intelligence (AI) has evolved enormously in the clinical context, or medical field [9,10]. In the medical field, Machine learning (ML) and Deep Learning (DL) algorithms prove their worth in implementing smart and automated AI-based systems [10–13]. Researchers have pulled their strings to develop more advanced frameworks that can be applied in various image-based applications. Convolutional Neural Network (CNN) is considered the state-of-the-art method in the analysis of visual imagery. In medical image analysis such as X-ray images, MRI images, CNN model and its derivations such as ResNet, VGG-16, GoogleNet, AlexNet, etc., have shown significant results in detection, recognition, and classification tasks [14]. However, deep learning architectures like CNN require immense computation resources as well as a lot of image data to train the proposed model [15]. Due to the lack of sufficient data and resources, the field of medical image analysis for skin diseases is yet to explore to the full extent. Pretrained CNN models have come to a point by researchers to aid the purpose. Besides, image analysis techniques such as Augmentation are widely used to construct a generalized model and robust systems where training data is inadequate.

CNN architectures like MobileNet and Xception are helping researchers to bring out new intelligent systems nowadays. For example, the MobileNet model shows high accuracy for the classification task [16] where welding defects from images were analyzed. In medical imaging, such as children's colonoscopy [17] combination of MobileNet with DenseNet is proposed for better classification results. In [18], lung diseases were analyzed and detected from chest X-ray images using the MobileNet model. In language processing tasks [19,20] MobileNet model was studied for the recognition task of Bangla characters which are handwritten and complex sign language translations. The Xception model is also widely used for different computer vision-based tasks. For example, chest X-ray images were analyzed using the Xception model in [21,22] to differentiate between COVID-19 lung condition and normal pneumonia. In [23], Xception based framework is used to classify and authenticate forensic images. Researchers also implemented this model for the garbage image classification task in [24] for the productive garbage management system.

1.3. Contribution

In this work, we implement an automated system based on computer vision-based techniques where two structured Convolutional Neural Network architectures MobileNet [25] and Xception [26], contribute to the recognition of different types of dermatological diseases, namely Atopic dermatitis, Eczema, Herpes, Nevus, and Melanoma. In order to construct an accurate model, we combined these two architectures with transfer learning and a real-time image augmentation process. In addition to that, we evaluated the effectiveness of our propositions by comparing the performance with state-of-the-art deep learning models such as ResNet50, InceptionV3, Inception-ResNet, and DenseNet.

Besides, we proposed and implemented a web-based architecture for the real-time recognition of diseases. We deployed our trained models on the web using Flask framework [27], and the recognition of skin diseases can be done remotely using this system. Our proposed approach can aid health professionals by recognizing different skin diseases more efficiently and making the diagnosis process more user-friendly for the patients. Moreover, besides pandemics and natural disasters, a cloud-based healthcare system can be built to operate the healthcare system remotely. Here we sum up the whole concept of this work's contribution below:

- Propose an automated framework for skin disease recognition based on pre-trained CNN architectures, namely MobileNet and Xception.
- For a more robust and generalization model, augmentation and transfer learning techniques are included.
- Propose and implement a web-based application to recognize skin diseases remotely.
- Evaluate the model's performance by comparing it with other deep learning models such as ResNet50, InceptionV3, Inception-ResNet, and DenseNet.

2. Literature review

Researchers were trying to develop an efficient and effective system that visually recognizes different classes of skin diseases. Some of the approaches include image processing techniques with statistical methods, texture, and color analysis. AD Mengistu, DM Alemayehu [4] proposed image processing techniques for recognizing and predicting skin cancers. Predefined classes of skin cancers collected from the American cancer society and DERMOFIT were used in this experiment. A hybrid method that integrates two image processing techniques, namely a Self-organizing map (SOM) and radial basis function (RBF), was used in this recognition task, and image features such as color, textures, and image structure were combined. Further, the acquired results were compared with other approaches such as KNN, Naïve Bayes, and ANN. The reported result revealed that the overall accuracy for this applied hybrid method was 93.15%. Manish Pawar et al. [5] Identify different skin disease conditions based on feed-forward backpropagation neural networks. Texture features were used as key attributes for image recognition purposes that were analyzed from the GLCM method. Three skin conditions were selected for the classification task, and the overall accuracy was reported at 66.66%. To enhance the scope for identifying multiple skin diseases, Li-sheng et al. [6] proposed a method that combines both color and texture features. The preprocessing task included noise and background removal through filtering and transformations. The GLCM approach was implemented to extract texture features such as contrast, correlation, entropy, etc., and for color feature extraction watershed algorithm was used. For this research purpose, three types of common skin diseases, namely herpes, dermatitis, and psoriasis, were classified using a support vector machine (SVM) classifier. The average accuracy while recognizing those 3 classes of skin disease images reached 90% using SVM classifier and combining color and texture features. Md. Nazrul Islam et al. [7] established a system for recognizing multiclass skin diseases that relied on image texture. Different preprocessing operations, such as resize, grayscale conversion, contrast enhancement, and noise removal were conducted for this experiment. Images textures were extracted using the GLCM method, and segmentation was carried out using Maximum Entropy Thresholding. Finally, the Backpropagation (BPN) algorithm was used to classify 3 different classes of skin disease images Eczema, Impetigo, and Psoriasis. The obtained accuracy for this method was reported at 80% along with sensitivity and specificity of 71.4% and 87.5%, respectively.

Rahat Yasir et al. [28] proposed a computer vision-based approach for recognizing skin diseases from images. Different preprocessing algorithms, like sharpening, median, smooth filter, binary mask, histogram, YCbCr, etc., were used for feature extraction. An artificial neural network (ANN) was used for training and test purposes. On a real-time dataset, the proposed model obtained a classification accuracy of 90%. To make a classification between skin conditions such as normal, spots, and wrinkles, Jhan S. Alarifi et al. [29] used traditional ML approaches based on SVM and CNN. SVM used feature extraction techniques like LPB and HOG. For CNN, GoogleNet architecture was implemented with different optimizers. The experimental result showed that GoogleNet with NAG optimizer outperformed SVM in all aspects, reached to an accuracy level of 89%. Yuexiang Li and Linlin Shen [30] proposed

deep learning methods for tasks like segmentation, extraction, and classification, which are involved in skin lesion detection. A fully convolutional residual network (FCRN) framework with a lesion index calculation unit (LICU) was used for segmentation and classification. On the other hand, feature extraction was carried out using the Lesion Feature Network (LFN) framework. The experiment was evaluated on ISIC 2017 dataset containing 2000 images for training and 150 images for validation. The accuracies of the proposed approaches for segmentation and classification were 75% and 91%, respectively, and the feature extraction task achieved an accuracy of 84%. An automated image recognition system was proposed by Jainesh Rathod et al. [31] for detecting skin diseases. Noise removal and image enhancement filters were used in the preprocessing phase. CNN algorithm was applied to the DermNet dataset as a feature extractor and classifier. An accuracy of 70% was reported in this experiment.

Min Chen et al. [32] proposed a real-time and dynamic framework for skin disease recognition that is composed of self-learning with a wide collection of data for effective interaction between users. A data filter algorithm was employed for the removal of unwanted data and feeding the network with valuable data. Information entropy was used for this filtering task. Three CNN learning models, namely LeNet-5, AlexNet, and VGG-16, were used for the classification and prediction task. The authors also test the reliability and validity of the proposed system by analyzing the computation and transmission delays of the system. This analysis, it had been shown that the communication delay of the AlexNet and LeNet models is 75 ms and 63 ms, respectively. Md Ashraful Alam Milton [33] experimented with Melanoma detection techniques where different deep neural networks like PNASNet-5-Large, InceptionResNetV2, SENet154, InceptionV4 were used. Images from ISIC 2018 dataset were used to train and test the proposed models. All the images were preprocessed by several operations such as normalizations, and augmentation before training. Parameters were initialized using the ImageNet model, and models were fine-tuned. The highest validation score was 76% which was reported for the PNASNet-5-Large model. For the construction of an automated computerized diagnosis system, Haofu Liao [34] proposed a method based on deep CNN. In this study, advanced CNN architecture such as VGG-16, VGG-19, and GoogleNet was implemented. The experiment was conducted on two different datasets, namely Dermnet and OLE, and the performance of the models was compared. All the models that were used in this study were pretrained on the ImageNet dataset. On the DermNet dataset, the top-5 accuracy was 91% using the VGG-16 model, while for the OLE dataset, the top-5 accuracy reached 69.5%. Shanthi et al. [35] suggested a method that was used to detect four types of skin diseases from the DermNet Dataset. The CNN architecture called ALexNet, which is utilized by 11 layers, was used for the detection task. The maximum pooling layer with a learning rate of 0.01 was used for the model training purpose. The highest accuracy was 93.3% for the Eczema herpeticum class.

Srinivasu et al. [36] combined MobileNet V2 with LSTM to classify skin disease. HAM10000 dataset was used in this experiment and the reported accuracy was 85%. They have also proposed a web application for the classification of skin decreases. In another work, Iqbal et al. [37] proposed multi-class classification for skin diseases using a deep CNN model and used 3 different datasets, namely ISIC-17, ISIC-18, and ISIC-19. The proposed experiment has achieved a specificity of 91% with the ISIC-17 dataset. Reis et al. [38] presented a CNN network based on inception block (InSiNet) to detect lesions. The proposed method reported an accuracy of 94.59% with the ISIC 2018 dataset. A details comparison between different CNN architectures was demonstrated in this study and InSiNet outperforms all of them. To classify skin cancer, Gupta et al. [39] used the CNN model that can work on both dermoscopic and photographic images. This approach obtained a classification accuracy of 83.2% using the Inception V3 model.

Kalaiyarivu and Nalini [40] studied a CNN-based approach to detect skin diseases by extracting color and texture features. The proposed CNN model achieved an accuracy of 87.5%. In another work, Kousis et al. [41] used 11 different CNN models to recognize skin cancer. In this approach, they have used the HAM10000 dataset, and they reported an accuracy of 92.25% using the DenseNet169 model. Ahmad el at. [42] proposed a hybrid classification approach using CNN and stacked BLSTM. In this work, BLSTM is used for feature extraction and then ensembled with a deep CNN network for the classification task. The authors experimented on two different datasets, one customized and another one HAM10000, and reported an accuracy of 91.73% and 89.47%, respectively. Aijaz et al. [43] proposed a deep learning-based application where different categories of skin diseases are classified. two different deep learning models CNN and LSTM were used in this approach. For better results, different pre-processing techniques such as augmentation, enhancement, and segmentation were employed in this study and achieved an accuracy of 84.2% and 72.3% for CNN and LSTM, respectively.

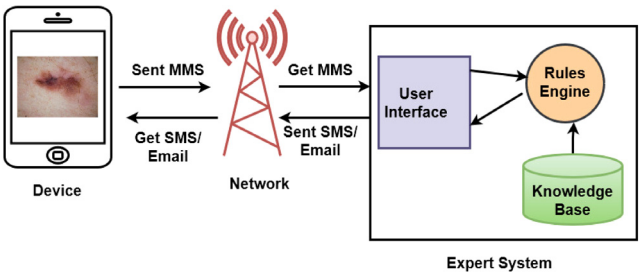


Fig. 1. Architecture of the expert system to recognize skin diseases.

3. System architecture and research methodology

This section scrutinizes the pertinent technologies and architectures that will be used to develop an automated system for different skin disease recognition.

3.1. System architecture

We presented a web-based medical expert system using deep learning framework for recognizing different skin diseases. The proposed overview of the system is illustrated in Fig. 1. The hypothesis behind this architecture is that a user captures an image of the diseased area using a smart device where the proposed application will be pre-installed. Then the image will be sent to the expert system through the application. Then feedback will be generated based on our trained model or expert system. The feedback will be returned to the user who seeks to identify or diagnose skin diseases via email or SMS.

3.2. Research methodology

Since we have minimal images to train our deep learning model, we propose models with real-time data augmentation and transfer learning approach integration. First, the acquisition of pretrained weights from tasks conducted on the ImageNet dataset is made in the building phase. Features from the training data are extracted as well as the labels in this phase. Second, a feature tensor combines these features according to the class labels. Third, the transfer learning approach uses pretrained weights that are acquired from a large dataset is applied. The domain knowledge from this phase is then transferred to the MobileNet and Xception model in the building phase. An unlabeled image data is then fed to the learned network. Finally, the model generates the class labels based on the knowledge it has gained from the previous phase. The schematic representation of our approach is given in Fig. 2. The required steps with the setup that will be carried out throughout the experiment are illustrated in Table 1.

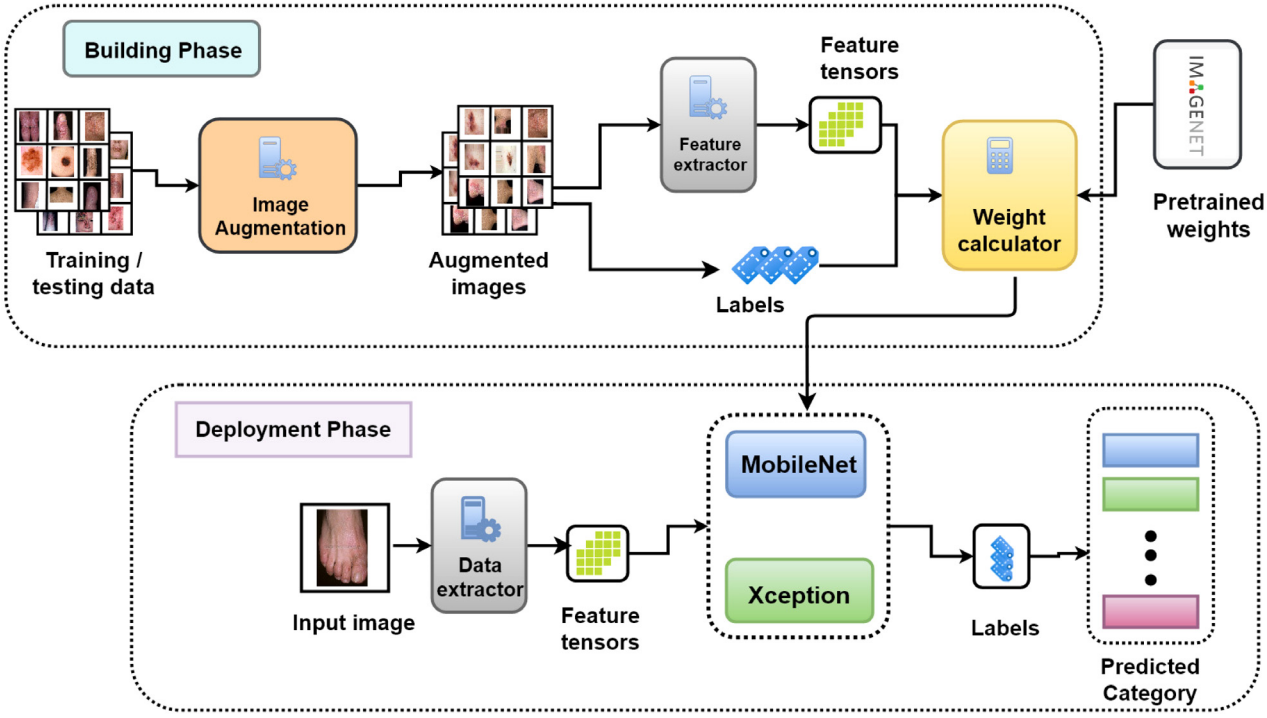


Fig. 2. Proposed system architecture. (A systematic representation of our proposed approach including data acquisition, preprocessing using augmentation, transfer learning, training, testing, and predictions carried out in building and deployment phases.)

Table 1
All the necessary steps with the setup that will be carried out throughout the experiment.

Algorithm: Experimental setup		
Input	*1.*	*Collected images of 5 classes of skin diseases .*
Environment	*2.*	*Google Colab.*
Configuration	*3.* *4.*	*Import all necessary libraries and packages .* *Import the images.*
Directories Configuration	*5.*	*Construct directories for Training, testing and validation.*
Training and Testing	*6.* *7.*	*Build CNN models. For transfer learning use a model trained on ImageNet dataset.* *Fine-tune the models by adding additional global average pooling layers, a fully connected layer and Softmax class*
Model Compilation	*8.* *9.* *10.* *11.*	*Model compile with RMSProp optimizer and learning rate of 0.001.* *Set 100 epochs for model fitting.* *Use val_accuracy monitor as model checkpoint.* *Save model.*
Performance Evaluation	*12.* *13.* *14.*	*Generate classification report and confusion matrix.* *Generate AUC–ROC curve.* *Generate model accuracy and loss reports.*
Prediction	*15.* *16.* *17.*	*Load best model.* *Load random images.* *Predict disease classes.*

We have implemented six different CNN-based architectures namely ResNet50, InceptionV3, Inception-ResNet, DenseNet, MobileNet, and Xception. But we focused more specifically on MobileNet and Xception Model. The remaining models are used in this study to compare the performance of our propositions.

3.2.1. Convolutional Neural Networks (CNN)

CNN is the most popular artificial neural network specially designed for computer vision-based applications that incorporates analyzing visual imagery [44]. The network takes an image as input and processes the image for extracting different features and patterns from that input image. These features are also made distinguishable by the network. Both spatial and temporal characteristics are captured using CNN. These characteristics are used in differentiating different classes of images. The feature detection task is the backbone of the CNN model, which has been carried out using the feature extractor filter or Kernel.

The learning process of CNN constitutes convolutional layers, non-linear processing units, and layers for subsampling tasks [45]. The working of CNN implements a layered architecture and presented in Fig. 3. Three main layers, namely convolution, pooling, and fully connected layer, are used to build a CNN model [46]. Convolutional layers have a convolutional kernel that works as a feature extractor. These kernels slice the input image into receptive fields. The relation between the input feature map and output feature map can be expressed using convolutional operation, i.e., $F(x,y) = (f * k)(x,y) = \sum_i \sum_j f(i,j)k(x-i, y-j)$. $F(x,y)$ and $f(x,y)$ corresponds to the output and input feature map, and k(x,y) represents the element of the corresponding kernel. The pooling layer involves an operation that sums up all the relevant and similar information from the neighborhood. The size of the input feature map has been reduced by cutting down the number of parameters. The pooling operation can be formulated using the equation, $Z = g_p(f)$ where Z is the polled feature map operating with input

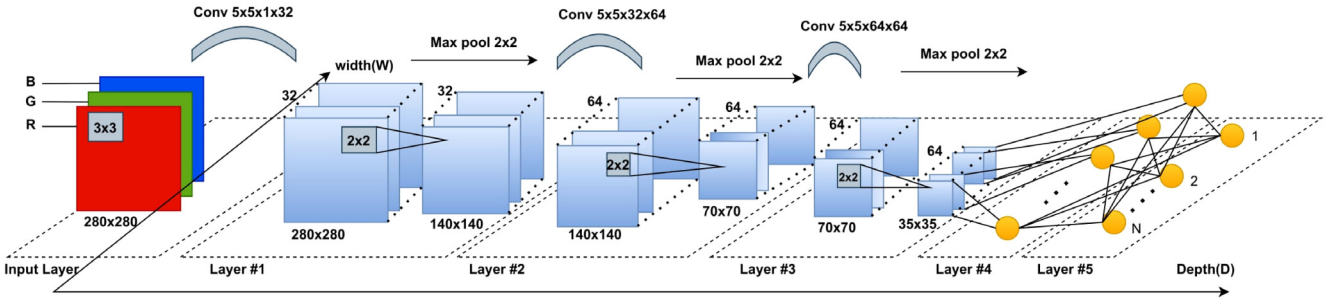


Fig. 3. A convolutional neural network (CNN) architecture with its dimensions. (A layered representation of CNN architecture for performing different operations like convolution, pooling, and consist of convolution layers, pooling layers and fully connected layer.)

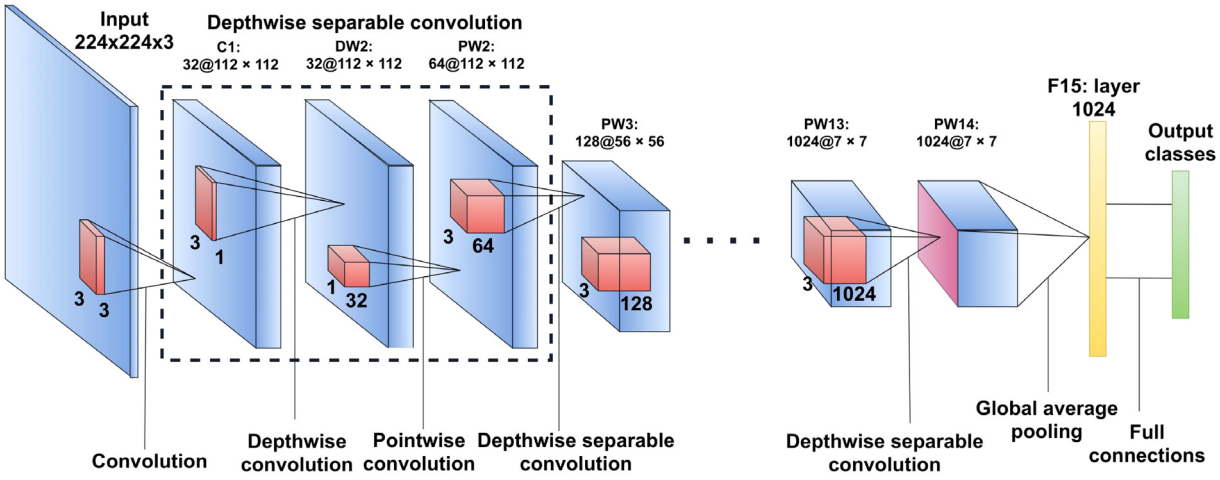


Fig. 4. Architecture of MobileNet. (A CNN architecture performing Depthwise and Pointwise convolution on the input image for the completion of the filtering task and the creation of linear output combinations.)

feature map f. Finally, the classification task has been carried out using a global operation carried out in a fully connected layer (FC). All the extracted features are analyzed in this layer and create a non-linearity between them.

3.2.2. MobileNet

MobileNet is a popular Deep CNN network, widely used in computer vision-based applications such as image classification, categorization or segmentation, etc., for its lightweight and small architecture and fast operational characteristics [25]. The fabrication of MobileNet is established on depthwise separable filters represented in Fig. 4. The main focus of this model is to optimize latency with a small network and make a model that is suitable for deploying on mobile devices. MobileNet architecture is incorporated with two steps, namely depthwise convolutions and pointwise convolutions. First, the feature extraction process is carried out by depthwise convolutions, where only a filter processes each input channel. Then the pointwise 1×1 convolution is applied that combines features obtained from depthwise convolutions. In depthwise separable convolutions, extraction of features, and combining those features are done by separate layers. This results in the reduction of computation time and computation cost, and model size.

There exist some architectural differences between the general convolutional layer and the depthwise convolutional layer. The input that is taken by a standard convolutional layer can be expressed as $D_F \times D_F \times M$ of feature map F and produces $D_G \times D_G \times N$ of feature map G. The value of $D_F \times D_F$ represents the dimension (height*width) of the input image and $D_G \times D_G$ represents the dimension (height*width) of the output image. Here N is the number of input channels or input depth, and M is the number of output channels or output depth. For standard convolution layers where K is the kernel of size $D_K \times D_K \times M \times N$ where $D_K \times D_K$ denotes the dimension of the kernel. The output feature map is given by the following equation

$$G_{(k,l,n)} = \sum_{(i,j,m)} K_{(i,j,m,n)} \cdot F_{(k+i-1,i+j-1,m)} \tag{1}$$

For the depthwise convolution layer, is depthwise convolution kernel is denoted by $\hat{K}$, and the size of this kernel can be computed as $D_K \times D_K \times M$. So the depthwise convolution for input depth can be written as

$$\hat{G}_{(k,l,m)} = \sum_{(i,j)} \hat{K}_{(i,j,m)} \cdot F_{(k+i-1,i+j-1,m)} \tag{2}$$

Here m_{th} filter in $\hat{K}$ applied to the m_{th} channel in F to produce the m_{th} channel of $\hat{G}$. The total computational cost for depthwise convolutions is given by $D_K.D_K.M.D_F.D_F$

3.2.3. Xception

Xception is another class of Deep CNN which is adapted from the Inception-V3 model [26]. The model is constructed based on the intuition of the depthwise separable convolutional module. Modification is made in the inception block of the Inception-V3 model. The modified architecture for Xception has a wider inception block than Inception-V3. It has spatial dimensions of 1×1, 5×5, and 3×3, which is replaced in the Xception model with a single dimension of size 3×3 and 1×1, i.e., Convolution part is divided into spatial and pointwise convolution. Fig. 5 illustrates the architecture of the Xception network. Firstly, 1×1 pointwise convolution is applied, and then a 3×3 depthwise convolution is applied [45]. This approach results in the

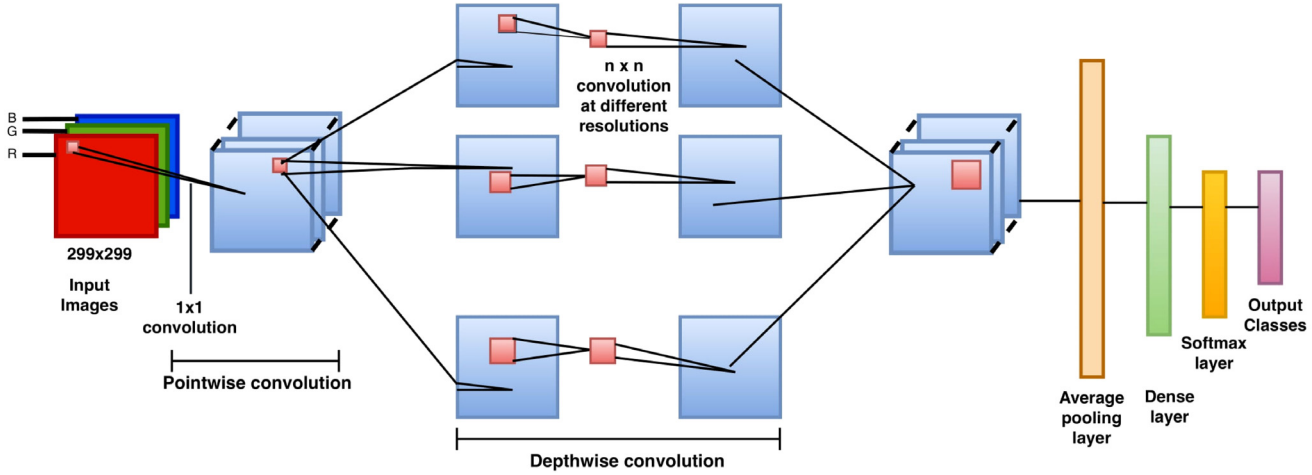


Fig. 5. Architecture of Xception. (A layered architecture of Xception consisting of 36 convolutional layers and 14 modules. It implements a 1 × 1 pointwise convolution followed by a 3 × 3 depth-wise convolution.)

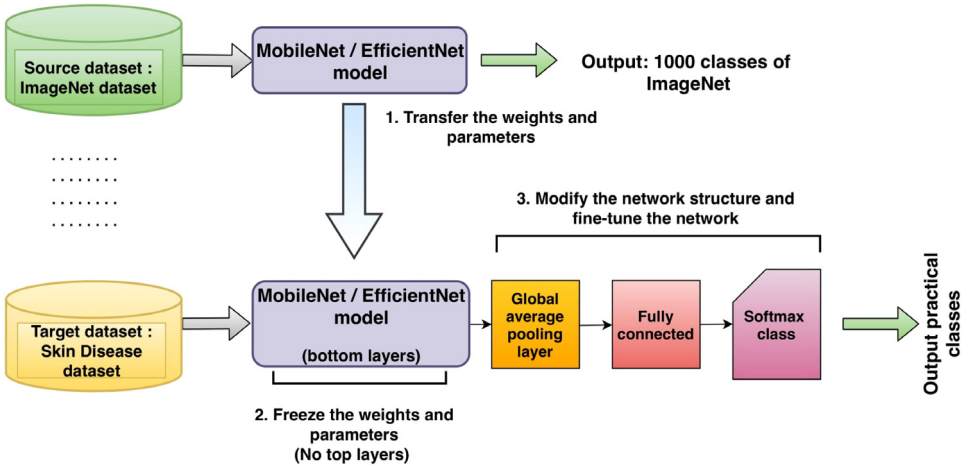


Fig. 6. The process of transfer learning. (Pretrained weights from earlier tasks conducted on a very large dataset has been used for the purpose of transporting knowledge. An additional global average pooling layer, a fully connected layer, and a Softmax layer are added for fine-tuning the network.)

reduction of parameters and layers and makes the network lightweight. Disengagement of this correlation is followed by Eqs. (3) and (4).

$$f^k_{(l+1)}(p,q) = \sum_{(x,y)} f^k_l(x,y).e^k_l(u,v) \tag{3}$$

$$F^k_{(l+2)} = g_c(F^k_{(l+1)}, K_{(l+1)}) \tag{4}$$

Here, F corresponds to the feature map of l transformation layers, (x, y) and (u, v) show the spatial indices of feature map F and kernel K having depth one. Kernel K is spatially convolved across feature map F. Here $g_c(.)$ indicates the convoluted operation. In total, a basic Xception model has 36 convolutional layers and 14 modules. Among these, 12 modules are connected with a residual layer boosting the merging process and paving the way for achieving higher accuracy. Architecturally, the Xception network consists of 3 flows, namely Entry flow, Middle flow, and Exit flow. Downsampling of input images with dimensionality reduction is carried out using the Entry flow part. Learning from features and optimizing those features is done by the Middle flow part of the network. Finally, the Exit flow carries out the integration of features.

3.3. Transfer learning

The Transfer Learning (TL) approach in the context of deep learning is a pervasive method in computer vision-related tasks. However, creating a robust and generalized deep learning model, it is highly required a lot of images and computational resources [47]. To overcome this, deep learning models can utilize the TL approach, in which a model that has been trained for one task can be used as a baseline model for another. This method of reusing models that have been trained previously with a large amount of data can be used in another training process that has a small amount of data and paves the way for achieving higher accuracy [48]. In general, weights are initialized using random numbers in the training process of neural networks. These assigned weights are then slowly updated during the training process. So in most cases, training with a small number of training data cannot achieve sufficient accuracy. To perform the transfer learning process, we should prepare a neural network model trained with many data that can handle similar types of data, which becomes the source model for transfer learning.

In the transfer learning process, features learned from huge image sets such as ImageNet are highly transferable to a variety of image recognition tasks [49]. This process is depicted in Fig. 6. Several ways to transfer knowledge from one model to another. One approach is to train the top layer of the already pretrained model and then replace it with a randomly initialized one. After that, the top layer parameters are trained for the new task while all other parameters remain fixed. This approach best suits a task where there is a maximum similarity between the pretrained model and the new task. If we have more data, then we can train the entire network by unfreezing these transferred parameters. Only the initial values of the parameters are transferred while

(a) Nevus (b) Melanoma (c) Herpes (d) Eczema (e) Atopic dermatitis

Fig. 7. Different skin diseases used in our approach.

Table 2
Overall dataset splitting.

Skin diseases	Training images	Validation images	Test images
Atopic dermatitis	2610	432	100
Eczema	4750	950	100
Herpes	4200	840	100
Nevus	1955	391	100
Melanoma	1720	344	100
Total	15235	2957	500

weights are initialized using pretrained models instead of initializing them randomly, boosting the convergence process.

4. Experimental evaluation and result analysis

4.1. Environment specifications

Image analysis or classification requires intense computing powers, and GPU (Graphics Processing Unit) can provide such computing compatibility. But GPU installation is expensive and requires additional hardware to support the computing task. So we use the Google Colab[1] platform to train our model, which provides us with high-end GPU on the cloud. It comes with all the necessary packages which are used in the training process, so there is no burden of installing packages or extra storage [50]. Google Colab comes with NVIDIA K80 GPU, GPU memory of 12 GB, Up to 2.91 teraflops double-precision rendition, and disk space of 358 GB. These specs give an enormous computation environment to train Deep Learning models.

4.2. Dataset description

We have used 5 classes of skin diseases, namely Atopic dermatitis, Eczema, Herpes, Nevus, and Melanoma. Since there is no available dataset that contains images of all these classes, we prepared our dataset by collecting images from two different sources. We have collected images for Atopic dermatitis, Eczema, Nevus, and Herpes from Dermnet [51]. For Melanoma images, we have used the HAM10000 dataset [52]. A total number of 18692 images are used in our approach, split for training, validation, and testing purposes. A glimpse of images constituting our dataset is given in Fig. 7 Splitting the dataset into training and testing datasets depicts in Table 2.

4.3. Data preprocessing

The proposed CNN architecture MobileNet and Xception require very less preprocessing images as they extract features directly from images. MobileNet model requires an input shape of 224×224, and the Xception model requires images of dimension 229×299. So firstly, images are resized according to the measurement for each model. Since a robust model requires many images to train and validate the model, we performed real-time image augmentation in our study, expanding our training data virtually. The augmentation operations that are performed in this task are flip, shift, and zoom. Both vertical and horizontal flips are performed to reverse the pixel columns or rows. Shift operation moves all of the pixels unidirectional, and using zoom operation, images are randomly zoomed by adding new pixel values.

Augmentation techniques that are used in our approach are illustrated in Fig. 8.

4.4. Performance evaluation metrics

The performance of a classifier is described through the confusion matrix, which gives an insight into the correct and incorrect predictions made by the classifiers [53]. A classifier is used to predict some classes that can be either true or false. There can be four cases as output while classifying some data belonging to more than one class. Firstly all the predictions (true or false) are correct, which is indicated by True Positive (TP) and True Negative (TN). However, there can be another case in which the prediction is true, but in reality, it is false, and vice-versa. These two cases are called False Positive (FP) and False Negative (FN). Not only that, we can calculate some more specific metrics from the confusion matrix that can be deciding factors for revealing the classification performance of our models. These metrics are Accuracy, Precision, Recall, and F1-score. These metrics are calculated using the following formulas.

Accuracy: Accuracy is the indicator of how well a model can predict true and false classes precisely and expressed using formula (5).

$$Accuracy = \frac{\sum_i^N M_i}{\sum_i^N |T_i|} \times 100\% \tag{5}$$

where, $\sum_i^N M_i$ indicates the total number of correct predictions, and $\sum_i^N |T_i|$ is the total number of predictions.

When it comes to binary classification, Accuracy is represented using the following formula (6)

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

where, TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives.

Precision: Precision indicates how well a classier performs in terms of predicting correct outcomes that are positive. Mathematically representation can be established using the formula (7)

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

Recall: Recall indicates the performance of a classier by measuring the proportion of true positive observations that were correctly predicted. Formally Eq. (8) defines Recall,

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

F1 score (F-measure): F1 score is the symphonic average of precision and recall. Formally it is represented mathematically as Eq. (9)

$$F1\ score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{9}$$

[1] https://colab.research.google.com/

(a) horizontal flip (b) width shift (c) height shift (d) zoom (e) vertical flip

Fig. 8. Augmentation techniques that are used in our approach.

Table 3
Class-wise classification results of MobileNet and Xception. (Values of evaluation metrics Precision, Recall, and F1-score for MobileNet and Xception model with **Transfer Learning** approaches is presented for each disease classes.)

Model	Class	Recall (%)	Precision (%)	F1 (%)
MobileNet	Atopic dermatitis	97.00	90.70	93.71
	Eczema	89.00	95.70	92.22
	Herpes	95.00	96.94	95.96
	Melanoma	100.00	97.08	98.51
	Nevus	99.00	100.00	99.50
Xception	Atopic dermatitis	96.00	97.00	96.50
	Eczema	90.00	95.74	92.80
	Herpes	99.00	92.52	95.65
	Melanoma	100.00	100.00	100.00
	Nevus	100.00	100.00	100.00

Table 4
Class-wise classification results of MobileNet and Xception. (Values of evaluation metrics Precision, Recall, and F1-score for MobileNet and Xception model with **without Transfer Learning and without augmentation** approaches is presented for each disease classes.)

Method	Class	Recall (%)	Precision (%)	F1 (%)
MobileNet	Atopic dermatitis	88.3	91.0	89.6
	Eczema	85.7	66.0	74.5
	Herpes	84.6	99.0	91.2
	Melanoma	89.4	93.0	91.1
	Nevus	100.	99.0	99.4
Xception	Atopic dermatitis	84.2	91.0	87.4
	Eczema	91.0	71.0	79.7
	Herpes	80.4	99.0	88.7
	Melanoma	97.8	93.0	95.3
	Nevus	100.0	96.0	97.9

Sometimes accuracy and F1-score are not enough for evaluating predictive models. So another metric which is called the Receiver Operating Characteristics curve or ROC curve, is also used for evaluation. With AUC, an accumulated measure of performance can be defined at every possible classification threshold. From the ROC curve, the area under the ROC curve (AUC) is induced, which is a compatibility indicator of a predictive model. Derivation of ROC can be done when the True Positive Rate (TPR) is plotted against False Positive Rate (FPR). True positive rate is nothing but Recall, and FPR is defined by an Eq. (10)

$$FPR = \frac{FP}{FP + TN} \tag{10}$$

4.5. Results

In this segment, we demonstrate the results of our proposed architectures (MobileNet and Xception) to scrutinize the robustness of the models. Additionally, the experiment was conducted on other deep learning models, ResNet50, InceptionV3, Inception-ResNet, and DenseNet, to compare and evaluate the performance of our propositions. Finally, we present the performance comparison of the proposed architectures with some graphical presentations and tables.

4.5.1. Classification performance of proposed MobileNet and Xception models

Classification results of our proposed models MobileNet and Xception according to our classes (skin disease) are illustrated in Tables 3 and 4. We have shown the results based on propositions transfer learning (TL) with augmentation for each model and without TL and augmentation. To give an overall insight into our classification results in terms of the number of right classifications and misclassification, We presented confusion matrices for MobileNet and Xception model in Fig. 9. Fig. 9(a) illustrates the produced confusion matrix of MobileNet architectures. From this representation, it can be observed that Herpes and Eczema classes have achieved 100% right prediction scores for this approach. The classification performance of Xception architecture is illustrated in Fig. 9(b).

A more comprehensive representation of classification results is depicted in Table 5 of our proposed MobileNet and Xception models.

Table 5
Overall classification report. (Comparison results between ResNet50, InceptionV3,Inception-ResNet, DenseNet, MobileNet, and Xception model based on average values of Precision, Recall, and F1-score.)

Model	Recall(%)	Precision(%)	F1(%)
ResNet50	87.00	87.00	87.00
Inception-V3	93.00	93.00	93.00
Inception-ResNet	95.00	95.00	95.00
DenseNet	93.00	93.00	93.00
MobileNet	96.00	96.00	96.00
Xception	97.00	97.00	97.00

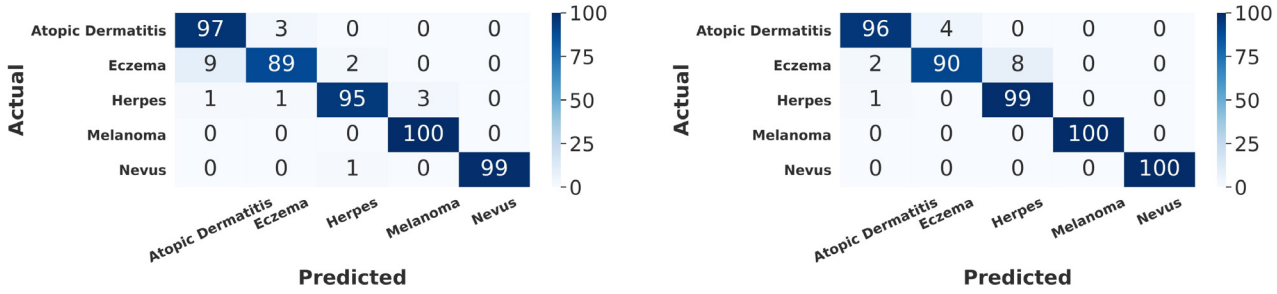
In addition, a comparison is established with the other models, such as ResNet50, InceptionV3, Inception-ResNet, and DenseNet.

Another compatibility indicator for our proposed models is ROC which is presented in Fig. 10. The highest reported micro average AUC score is 0.9974, which is reported for the MobileNet model. The lowest micro AUC score is reported for the ResNet50 model. The ROC of the Xception model is the second highest, which is 0.9972. Other models also showed good AUC scores.

4.5.2. Prediction accuracy and loss

In this segment, accuracy and loss for our approaches are depicted for all six models. In Table 6, validation and testing accuracy and loss are presented in terms. The highest testing accuracy is 97.00%, and the lowest loss is 0.16, reported for the Xception model with TL and augmentation. For MobileNet, the highest accuracy is 96.00%. ResNet50 showed the lowest test accuracy (86.60%) and highest loss score (2.40) compared to other models.

In Fig. 11(a) and (b), a line chart is illustrated for accuracy and loss for the MobileNet and Xception model for 100 epochs. It is seen from Fig. 11(a) that accuracy is pretty high and consistent for the approach using TL and augmentation for the MobileNet model. There exist some reductions and fluctuations per epoch for both models. For loss, Fig. 11(a) demonstrates that the lowest loss rate is gained from implementing both TL and augmentation approaches. For Xception Model, Fig. 11(b) demonstrates accuracy and loss for each epoch. Like MobileNet, here also observed high and consistent accuracy scores per epoch by implementing TL and augmentation. From Fig. 11(b), insight



(a) MobileNet (b) Xception

Fig. 9. Confusion matrix presenting total number of right and wrong prediction that occurs in the testing process for MobileNet and Xception model.

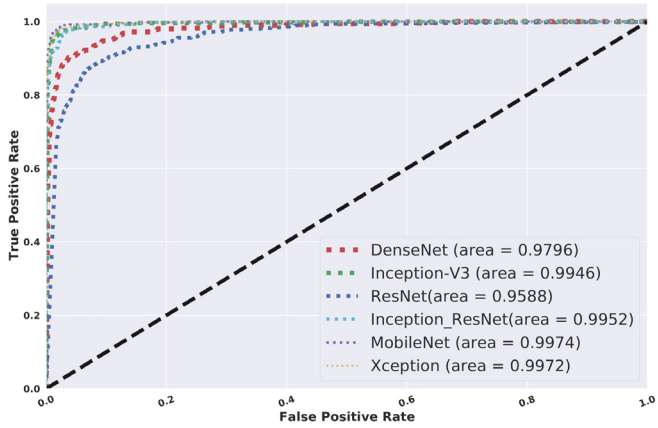


Fig. 10. ROC curve for deep learning models. This representation depicts the micro areas under the ROC curve (AUROC) for each of the models.

Table 6
Accuracy and Loss for the best models.

Model	Accuracy (%)		Loss	
	Validation	Test	Train	Validation
ResNet50	96.70	86.60	0.19	2.40
Inception-V3	98.0	93.0	0.05	0.45
Inception-ResNet	98.80	94.80	0.99	0.42
Densenet	94.20	92.80	0.98	0.34
MobileNet (Without TL)	94.45	89.38	0.8	1.77
Xception (Without TL)	95.55	89.79	0.76	1.13
MobileNet (Proposed with TL)	96.00	96.0	0.15	0.21
Xception (Proposed with TL)	97.94	97.0	0.07	0.16

into the loss per epoch can be achieved. Here low loss scores were reported per epoch by implementing Tl and augmentation.

Finally, the running time of our training process is given in Table 7 for each of our models. MobileNet model with TL and augmentation takes the shortest time (7869 s) to complete the 100 epochs. The longest time to complete the execution is reported for the Inception-ResNet model with the TL approach, 22971 s. The exception model also showed less time to complete the training process with 10877 s.

We have also presented a comparison between recent deep learning approaches that are proposed in different computer vision-based works in Table 8. From this comparison, it can be clearly derived that our models with augmentation and transfer learning techniques have better prediction accuracy.

The effectiveness of our approach in recognizing skin diseases is depicted in Figs. 12 and 13. we have used both MobileNet and Xception models with TL and augmentation to predict diseases as a part of deployment phases. From this presentation, it can be seen that Both

Table 7
Total running times for each model.

Model	Runtime (s)
ResNet50	14514 s
Inception-V3	5436 s
Inception-ResNet	22971 s
Densenet	16379
MobileNet	7869 s
Xception	10877 s

models can accurately predict the classes of the respective diseases. We can see one misclassification case for the MobileNet model. But the Xception model can successfully identify all skin diseases.

4.6. Result analysis

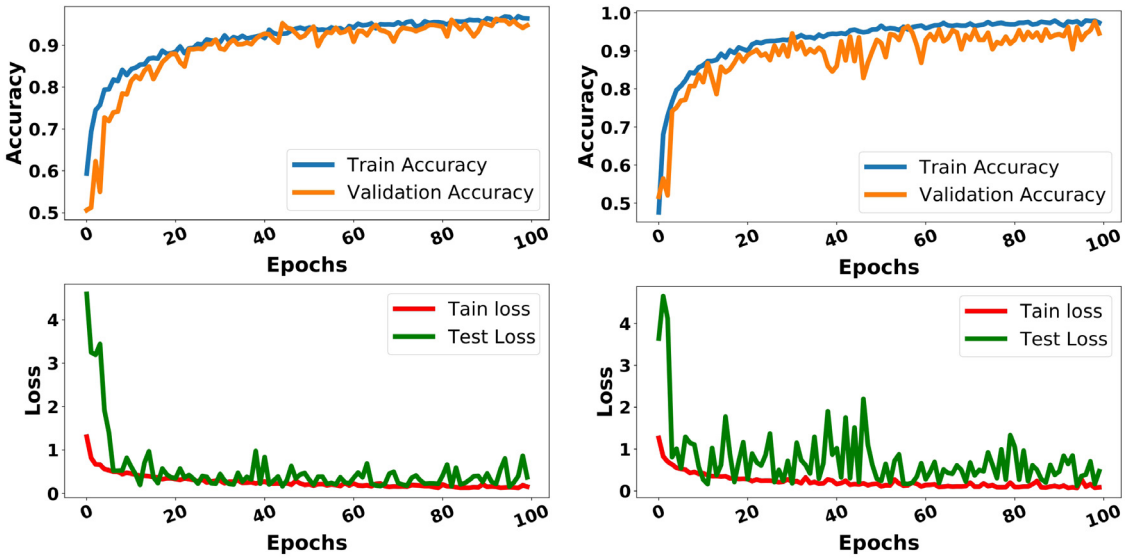
In this research work, we proposed implementing two deep learning-based architectures MobileNet and Xception, in recognizing different classes of skin diseases for computer vision-based applications. Besides, other deep learning models such as ResNet50, InceptionV3, Inception-ResNet, and DenseNet were also implemented to compare our approaches' effectiveness. Finally, we scrutinize the performance of our different propositions for skin recognition tasks based on classification reports, confusion matrix, ROC curves, and classification accuracy.

From Table 3, the decision can be reached about class-wise classification for both models. The highest Precision score is 100%, which is achieved for the Nevus class using both MobileNet and Xception models. Additionally Xception model also achieved a precision score of 100 for the Melanoma class. This tells us that our approaches result in a very good measure of the positive predictions that were actually correct. For Recall, a maximum of 100 scores is observed for nevus and Melanoma classes in Xception and melanoma classes in MobileNet. Since we have used an imbalanced dataset, F1-score can be a deciding factor. The maximum score is achieved for Melanoma and Nevus class using the Xception model. From Table 4, it can be seen that both Precision and Recall and F1-score are much lower for cases with no TL and augmentation. We observed the highest F1-score of 97% for the Xception (TL+A) model and 96.38% for MobileNet (TL+A) model. This is an indication that our proposed approach with TL and augmentation have good classification capability for imbalanced dataset.

The more comprehensive representation of Precision, Recall, and F1-score is given in Table 5, where an overall score for each of the metrics is given for each of the models. The highest precision is 97.05%, which is reported for the Xception model. This means that Xception models predict the correct class of skin disease most of the time. The highest recall value is 97.00% for Xception and 96.00% for the MobileNet model, i.e., both of these modes correctly identify most skin diseases. However, other models such as ResNet50 performed poorly, achieving a low score. We observed the highest F1-score of 97.00% for the Xception model and 96.38% for the MobileNet model.

Table 8
Comparison between existing approaches and our proposed approaches.

Method/Work done	Dataset	Used architecture	Classification accuracy	Best model
Yasir et al. [28]	775 clinical images	CNN with adaptive learning	90%	CNN
Alarifi et al. [29]	Clinical images	SVM + CNN	89%	CNN with SVM
Li and Shen [30]	ISIC 2017	FCRN with LICU	91%	FCRN
Rathod et al. [31]	DermNet	CNN	70%	CNN
Milton [33]	ISIC 2018	CNN (PNASNet-5-Large, InceptionResNetV2, SENet154, InceptionV4)	76%,70%, 74%, 67%	PNASNet-5-Large
Liao [34]	DermNet and OLE	CNN (VGG16)	91% (DermNet), 69.5% (OLE)	VGG16
Shanthi et al. [35]	DermNet	CNN (ALexNet)	93.3%	ALexNet
Kalaiyarivu and Nalini [40]	Clinical images	CNN	87.5%	CNN
Kousis et al. [41]	HAM10000	CNN	92.25%	DenseNet169
Ahmad et al. [42]	Customized	CNN + stacked BLSTM	91.73%	–
Gupta et al. [39]	ISIC	VGG16, VGG19, and Inception V3	82.4%, 83.0%, 83.2%	Inception V3
Proposed	DermNet + ISIC 2018	ResNet50, InceptionV3, Inception-ResNet , DenseNet, MobileNet, and Xception	86.60%, 93%, 94.80%, 92.80%, 96%, and 97%	**Xception**



(a) MobileNet (b) Xception

Fig. 11. Accuracy and Loss for MobileNet and Xception model with transfer learning and augmentation techniques.

Fig. 12. Predicting skin diseases using MobileNet model.

This indicates that our proposed approach with TL and augmentation has good classification capability for imbalanced datasets than other models presented in this study.

For illustrating the entire classification and misclassification, the confusion matrix as a heatmap is depicted in Fig. 9. Using the transfer learning and augmentation approach, both our models performed very satisfactorily, outperforming other models. MobileNet and Xception models reported only 20 and 15 misclassification cases, respectively. Accuracy and loss reported by our models are presented in Table 6. The highest classification accuracy is 97.00%, which is observed for the Xception model. The MobileNet model also gives a tremendous performance with a classification accuracy of 96.00%. ResNet seems to be a bad choice in terms of testing accuracy achieving 86.60% testing accuracy. Both models with TL and augmentation reported very low loss scores also. But approach with no TL and augmentation reported a higher loss score with low accuracy than other approaches.

With the ROC curve presented in Fig. 10, a relation is established between the false positive rate and the true positive rate. The highest

Fig. 13. Predicting skin diseases using Xception model.

Fig. 14. Developed web interface and application using a skin disease image.

ROC–AUC measures are obtained for TL and augmentation approaches. Finally, Table 7 presents the runtime measures. But when it comes to producing a more robust and accurate classifier, the runtime is a minor fact, whereas accuracy and other evaluation metrics are the major ones to consider.

4.7. Deployment of web application

Finally, we use the Flask [54] web framework to deploy our trained model. We created a web application that detects skin conditions by analyzing the skin photograph supplied by the client. To deploy the flask, we need two routes. First, we have created an index page route, which will help the users to upload their images. Finally, a predicted route will create an inference from our saved model.

The web application is created using the Xception model that has been trained using our skin dataset. Fig. 14 shows the developed web interface for the clients. The user uploads the image of a suspected diseased area using any smart device and submits it to the developed expert system or the web application through the interface. Then feedback will be generated based on our trained model by classifying the image based on different skin conditions.

5. Conclusion and future work

This paper has suggested a computer vision-based approach to recognize five skin diseases Atopic dermatitis, Eczema, Herpes, Nevus, and Melanoma. Two state-of-the-art deep learning models MobileNet and Xception, are implemented to create an automated system. We proposed approaches with transfer learning and augmentation techniques and evaluated each of the models in computer vision-based recognition tasks. Augmentation techniques pave the way for achieving a robust model by increasing the training data, whereas transfer learning enables reusing the weights from pretrained models. Integrating both transfer learning and augmentation techniques with MobileNet and Xception models proved to be a sophisticated approach to our disease recognition task, outperforming other deep learning models ResNet50, InceptionV3, Inception-ResNet, and DenseNet. The MobileNet model has achieved a classification accuracy of 96.00% and an F1-score of 96.38%. On the other hand, both test accuracy and F1-score of the Xception model reached 97.00%. In addition, We used the Flask web framework to deploy our trained model by creating a web application that detects skin conditions by analyzing the skin photograph supplied by the client. The presented approaches can help recognize and diagnose different dermatological diseases and aid health professionals in providing a better healthcare system. Finally, we have built a web-based framework to identify skin diseases.

For future studies, experiments will be carried out using a more diverse dataset. Only five classes of skin diseases are studied. So in the future, we plan to extend our experiment by adding more classes of skin diseases. Besides, the approach presented in this paper can be further enhanced by ensembling different deep learning models. Recently, transformer-based models such as Vision Transformers (ViTs) and MobileViT have been widely used in image processing tasks. So one of our future research directions could be to develop a transformer-based image recognition model for skin disease recognition. Our experiments took a huge computation time, and reducing the computation time in deep learning approaches could be another potential research direction of our work.

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.

Data availability

Data will be made available on request.

References

[1] V. Balaji, S. Suganthi, R. Rajadevi, V.K. Kumar, B.S. Balaji, S. Pandiyan, Skin disease detection and segmentation using dynamic graph cut algorithm and classification through Naive Bayes classifier, Measurement (2020) 107922.

[2] American Cancer Society, Cancer facts & figures for hispanics/latinos 2018–2020, 2020.

[3] R. Kasmi, K. Mokrani, Classification of malignant melanoma and benign skin lesions: implementation of automatic ABCD rule, IET Image Process. 10 (6) (2016) 448–455, http://dx.doi.org/10.1049/iet-ipr.2015.0385.

[4] A.D. Mengistu, D.M. Alemayehu, Computer vision for skin cancer diagnosis and recognition using RBF and SOM, Int. J. Image Process. (IJIP) 9 (6) (2015) 311–319.

[5] M. Pawar, D.K. Sharma, R. Giri, Multiclass skin disease classification using neural network, Int. J. Comput. Sci. Inform. Technol. Res. 2 (4) (2014) 189–193.

[6] L.-s. Wei, Q. Gan, T. Ji, Skin disease recognition method based on image color and texture features, Comput. Math. Methods Med. 2018 (2018).

[7] M.N. Islam, J. Gallardo-Alvarado, M. Abu, N.A. Salman, S.P. Rengan, S. Said, Skin disease recognition using texture analysis, in: 2017 IEEE 8th Control and System Graduate Research Colloquium, ICSGRC, 2017, pp. 144–148, http://dx.doi.org/10.1109/ICSGRC.2017.8070584.

[8] A. Nawar, N.K. Sabuz, S.M.T. Siddiquee, M. Rabbani, A.A. Biswas, A. Majumder, Skin disease recognition: A machine vision based approach, in: 2021 7th International Conference on Advanced Computing and Communication Systems, vol. 1, ICACCS, 2021, pp. 1029–1034, http://dx.doi.org/10.1109/ICACCS51430.2021.9441980.

[9] F. Curia, Features and explainable methods for cytokines analysis of Dry Eye Disease in HIV infected patients, Healthc. Anal. 1 (2021) 100001.

[10] V. Chang, V.R. Bhavani, A.Q. Xu, M. Hossain, An artificial intelligence model for heart disease detection using machine learning algorithms, Healthc. Anal. 2 (2022) 100016.

[11] S. Dev, H. Wang, C.S. Nwosu, N. Jain, B. Veeravalli, D. John, A predictive analytics approach for stroke prediction using machine learning and neural networks, Healthc. Anal. 2 (2022) 100032, http://dx.doi.org/10.1016/j.health.2022.100032, URL https://www.sciencedirect.com/science/article/pii/S2772442522000090.

[12] R. AlSaad, Q. Malluhi, I. Janahi, S. Boughorbel, Predicting emergency department utilization among children with asthma using deep learning models, Healthc. Anal. 2 (2022) 100050, http://dx.doi.org/10.1016/j.health.2022.100050, URL https://www.sciencedirect.com/science/article/pii/S2772442522000181.

[13] M. Ahammed, M.A. Mamun, M.S. Uddin, A machine learning approach for skin disease detection and classification using image segmentation, Healthc. Anal. 2 (2022) 100122, http://dx.doi.org/10.1016/j.health.2022.100122, URL https://www.sciencedirect.com/science/article/pii/S2772442522000624.

[14] S. Serte, A. Serener, F. Al-Turjman, Deep learning in medical imaging: A brief review, Trans. Emerg. Telecommun. Technol. (2020) e4080.

[15] N.C. Thompson, K. Greenewald, K. Lee, G.F. Manso, The computational limits of deep learning, 2020, arXiv preprint arXiv:2007.05558.

[16] H. Pan, Z. Pang, Y. Wang, Y. Wang, L. Chen, A new image recognition and classification method combining transfer learning algorithm and MobileNet model for welding defects, IEEE Access 8 (2020) 119951–119960.

[17] W. Wang, Y. Li, T. Zou, X. Wang, J. You, Y. Luo, A novel image classification approach via dense-MobileNet models, Mob. Inf. Syst. 2020 (2020).

[18] K. Sriporn, C.-F. Tsai, C.-E. Tsai, P. Wang, Analyzing lung disease using highly effective deep learning techniques, Healthcare 8 (2) (2020) 107.

[19] T. Ghosh, M.M.-H.-Z. Abedin, S.M. Chowdhury, Z. Tasnim, T. Karim, S.S. Reza, S. Saika, M.A. Yousuf, Bangla handwritten character recognition using MobileNet V1 architecture, Bullet. Electr. Eng. Inform. 9 (6) (2020) 2547–2554.

[20] T.M. Angona, A. Siamuzzaman Shaon, K.T.R. Niloy, T. Karim, Z. Tasnim, S. Reza, T.N. Mahbub, Automated bangla sign language translation system for alphabets by means of MobileNet, Telkomnika 18 (3) (2020).

[21] M. Rahimzadeh, A. Attar, A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of xception and ResNet50V2, Inform. Med. Unlocked (2020) 100360.

[22] E. Ayan, H.M. Ünver, Diagnosis of pneumonia from chest X-Ray images using deep learning, in: 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science, EBBT, IEEE, 2019, pp. 1–5.

[23] L. Yang, P. Yang, R. Ni, Y. Zhao, Xception-based general forensic method on small-size images, in: Advances in Intelligent Information Hiding and Multimedia Signal Processing, Springer, 2020, pp. 361–369.

[24] C. Shi, R. Xia, L. Wang, A novel multi-branch channel expansion network for garbage image classification, IEEE Access 8 (2020) 154436–154452.

[25] A.G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, MobileNets: Efficient convolutional neural networks for mobile vision applications, 2017, arXiv:1704.04861.

[26] F. Chollet, Xception: Deep learning with depthwise separable convolutions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 1251–1258.

[27] D.S. Reddy, P. Rajalakshmi, A novel web application framework for ubiquitous classification of fatty liver using ultrasound images, in: 2019 IEEE 5th World Forum on Internet of Things (WF-IoT), IEEE, 2019, pp. 502–506.

[28] R. Yasir, M.A. Rahman, N. Ahmed, Dermatological disease detection using image processing and artificial neural network, in: 8th International Conference on Electrical and Computer Engineering, IEEE, 2014, pp. 687–690.

[29] J.S. Alarifi, M. Goyal, A.K. Davison, D. Dancey, R. Khan, M.H. Yap, Facial skin classification using convolutional neural networks, in: International Conference Image Analysis and Recognition, vol. 10317, Springer, Cham, 2017, pp. 479–485, http://dx.doi.org/10.1007/978-3-319-59876-5_53.

[30] Y. Li, L. Shen, Skin lesion analysis towards melanoma detection using deep learning network, Sensors 18 (2) (2018) 556.

[31] J. Rathod, V. Wazhmode, A. Sodha, P. Bhavathankar, Diagnosis of skin diseases using convolutional neural networks, in: 2018 Second International Conference on Electronics, Communication and Aerospace Technology, ICECA, IEEE, 2018, pp. 1048–1051.

[32] M. Chen, P. Zhou, D. Wu, L. Hu, M.M. Hassan, A. Alamri, AI-skin: Skin disease recognition based on self-learning and wide data collection through a closed-loop framework, Inf. Fusion 54 (2020) 1–9.

[33] M.A.A. Milton, Automated skin lesion classification using ensemble of deep neural networks in isic 2018: Skin lesion analysis towards melanoma detection challenge, 2019, arXiv preprint arXiv:1901.10802.

[34] H. Liao, A deep learning approach to universal skin disease classification, 2015.

[35] T. Shanthi, R. Sabeenian, R. Anand, Automatic diagnosis of skin diseases using convolution neural network, Microprocess. Microsyst. (2020) 103074.

[36] P.N. Srinivasu, J.G. SivaSai, M.F. Ijaz, A.K. Bhoi, W. Kim, J.J. Kang, Classification of skin disease using deep learning neural networks with MobileNet V2 and LSTM, Sensors 21 (8) (2021) 2852.

[37] I. Iqbal, M. Younus, K. Walayat, M.U. Kakar, J. Ma, Automated multi-class classification of skin lesions through deep convolutional neural network with dermoscopic images, Comput. Med. Imaging Graph. 88 (2021) 101843, http://dx.doi.org/10.1016/j.compmedimag.2020.101843, URL https://www.sciencedirect.com/science/article/pii/S0895611120301385.

[38] H.C. Reis, V. Turk, K. Khoshelham, S. Kaya, InSiNet: a deep convolutional approach to skin cancer detection and segmentation, Med. Biol. Eng. Comput. 60 (3) (2022) 643–662.

[39] S. Gupta, A. Panwar, K. Mishra, Skin disease classification using dermoscopy images through deep feature learning models and machine learning classifiers, in: IEEE EUROCON 2021 - 19th International Conference on Smart Technologies, 2021, pp. 170–174, http://dx.doi.org/10.1109/EUROCON52738.2021.9535552.

[40] M. Kalaiyarivu, N. Nalini, Hand image based skin disease identification using machine learning and deep learning algorithms, ECS Trans. 107 (1) (2022) 17381.

[41] I. Kousis, I. Perikos, I. Hatzilygeroudis, M. Virvou, Deep learning methods for accurate skin cancer recognition and mobile application, Electronics 11 (9) (2022) 1294.

[42] B. Ahmad, M. Usama, T. Ahmad, S. Khatoon, C.M. Alam, An ensemble model of convolution and recurrent neural network for skin disease classification, Int. J. Imaging Syst. Technol. 32 (1) (2022) 218–229.

[43] S.F. Aijaz, S.J. Khan, F. Azim, C.S. Shakeel, U. Hassan, Deep learning application for effective classification of different types of psoriasis, J. Healthc. Eng. 2022 (2022).

[44] C.D.S. Duong, Automated fruit recognition using EfficientNet and MixNet, Comput. Electron. Agric. 171 (2020) 105326.

[45] A. Khan, A. Sohail, U. Zahoora, A.S. Qureshi, A survey of the recent architectures of deep convolutional neural networks, Artif. Intell. Rev. 53 (8) (2020) 5455–5516.

[46] Q. Li, W. Cai, X. Wang, Y. Zhou, D.D. Feng, M. Chen, Medical image classification with convolutional neural network, in: 2014 13th International Conference on Control Automation Robotics & Vision, ICARCV, IEEE, 2014, pp. 844–848.

[47] O. Ukwandu, H. Hindy, E. Ukwandu, An evaluation of lightweight deep learning techniques in medical imaging for high precision COVID-19 diagnostics, Healthc. Anal. 2 (2022) 100096, http://dx.doi.org/10.1016/j.health.2022.100096, URL https://www.sciencedirect.com/science/article/pii/S2772442522000417.

[48] K. Guzel, G. Bilgin, Classification of breast cancer images using ensembles of transfer learning, Sakarya Üniv. Bilimleri EnstitÜSÜ Dergisi 24 (5) (2020) 791–802.

[49] T.H. Sanford, L. Zhang, S.A. Harmon, J. Sackett, D. Yang, H. Roth, Z. Xu, D. Kesani, S. Mehralivand, R.H. Baroni, et al., Data augmentation and transfer learning to improve generalizability of an automated prostate segmentation model, Am. J. Roentgenol. (2020) 1–8.

[50] E. Bisong, Google colaboratory, in: Building Machine Learning and Deep Learning Models on Google Cloud Platform, Springer, 2019, pp. 59–64.

[51] Dermnet, 2020, URL http://www.dermnet.com/. (Accessed 04 November 2020).

[52] P. Tschandl, C. Rosendahl, H. Kittler, The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions, Sci. Data 5 (2018) 180161.

[53] N.D. Marom, L. Rokach, A. Shmilovici, Using the confusion matrix for improving ensemble classifiers, in: 2010 IEEE 26-Th Convention of Electrical and Electronics Engineers in Israel, IEEE, 2010, pp. 000555–000559.

[54] P. Singh, A. Verma, J.S.R. Alex, Disease and pest infection detection in coconut tree through deep learning techniques, Comput. Electron. Agric. 182 (2021) 105986.