IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 14, No. 6, December 2025, pp. 4943~4956
ISSN: 2252-8938, DOI: 10.11591/ijai.v14.i6.pp4943-4956

Melanoma classification using ensemble deep transfer learning

Soumya Gadag[1], Panduranga Rao Malode Vishwanatha[2], Virupaxi Balachandra Dalal[3]
[1]Department of Electronics and Communication, Jain College of Engineering and Research, Karnataka, India
[2]Department of Computer Science Engineering, FET-JAIN Deemed to be University, Bangalore, India
[3]Department of Electronics and Communication, Jain College of Engineering and Research, Belagavi, Karnataka, India

Article Info

Article history:

Received Apr 30, 2025
Revised Aug 11, 2025
Accepted Sep 7, 2025

Keywords:

Classification
Deep learning
Image processing
Melanoma
Transfer learning

ABSTRACT

Melanoma, a type of skin cancer, poses significant challenges in early detection and diagnosis. Several methods for early melanoma detection, including visual inspection and several machine learning models, face challenges with accuracy. To overcome these issues, deep learning has been widely adopted in various biomedical applications. In this work, we employ deep transfer learning methods to classify melanoma. Firstly, we collect publicly available datasets containing melanoma images, their corresponding ground truth for segmentation, and class labels. Subsequently, we perform data preprocessing, normalization, and label encoding to address issues of varied illumination, image noise, and data imbalance. Next, we conduct feature extraction utilizing the previously trained deep learning models, VGG, ResNet, InceptionResNet, and MobileNet. The characteristic vectors obtained from each model are fused to produce a comprehensive depiction among the provided pictures. In the classification stage, we employ ensemble learning using transfer learning models, including EfficientNet, Xception, and DenseNet. These models are trained on the final feature vector to classify melanoma images effectively. The effectiveness of the suggested method is verified using publicly available ISIC 2017–2020 datasets, these model reports average accuracy scores of 96.10%, 97.23%, 97.50%, 98.33%, and 98.60%, in that order.

Corresponding Author:

Soumya Gadag
Department of Electronics and Communication Engineering, Jain College of Engineering and Research
Karnataka, India
Email: soumya.gadag5@gmail.com

1. INTRODUCTION

Skin cancer is the most prevalent form of cancer, surpassing all other types combined in terms of diagnosis rates annually. In the United States alone, there are approximately 9,500 fresh cases identified daily, as reported by the Skin Cancer Foundation in 2017 [1]. By 2040, there will likely be close to half a million occurrences of skin cancer, with melanoma being the deadliest type, marking a staggering 62% surge since 2018. The severity of the situation is emphasized by the alarming fact that one person loses their life to skin cancer every four minutes, prompting dermatologists worldwide to classify its rising incidence as a global epidemic, as noted by Melanoma UK in 2020 [2].

Early detection and intervention, especially for melanoma, emerge as pivotal factors in improving the survival rates amid the mounting cases [3]. Excessive contact to ultraviolet (UV) radiation stands out as the primary identifiable culprit behind skin cancer [4]. Natural sunshine or other UV sources, such as indoor tanning machines, may be the source of this exposure (Cancer Research UK) [5]. Significantly, reduced ozone levels raise the risk of exposure to natural sunshine by increasing ground-level UV radiation (Department for Environment Food & Rural Affairs, 2020) [6]. Furthermore, regions situated closer to the

equator witness an uptick in non-melanoma skin cancer cases due to elevated UV radiation levels. Furthermore, lifestyle factors such as poor dietary choices, alcohol consumption, and smoking also contribute to the modifiable risk profile associated with skin cancer. Therefore, early detection and prevention of these cancers is studied widely.

Several methods have been introduced for melanoma detection; however, imaging-based methods have been widely adopted in various biomedical applications. Dermatologists can diagnose malignant lesions through dermoscopic visual examinations. The diverse textures and wounds present on the skin surface can make detecting skin cancer challenging using dermoscopy. Yet, accurately diagnosing skin cancer through manual examination of dermoscopic images is difficult. The accuracy of lesion diagnosis is heavily influenced by the dermatologist's experience. The only alternative techniques for detecting skin cancer that are now available are dermoscopy, biopsy, and macroscopic inspection. Because skin lesions are complex, they require more time and care. The dermatologists perform the observation with naked eyes, dermoscopy mechanisms and biopsy. Therefore, the accuracy of these systems relies on clinician's skill. A significant amount of research has been dedicated to developing computer-based image analysis algorithms for the early and rapid diagnosis of skin cancer, aiming to overcome the previously mentioned challenges, Primarily, these algorithms have been parametric, relying on normally distributed data for operation. However, due to the uncontrollable nature of the data, these methods often fail to provide accurate diagnoses.

Numerous medical imaging researchers have introduced computer-aided design (CAD) techniques. This four-step CAD process encompasses image pre-processing, identification of affected areas, feature extraction, and classification. Several methods have been introduced using computer vision approach such as artificial neural networks (ANN), decision trees (DT), and support vector machines (SVM). The research in [7], [8] provides a thorough analysis of various techniques. Nevertheless, there are a number of data processing limitations with machine learning techniques, including the need for better contrast, noise-free, and cleaned images. Moreover, a number of criteria, including structural traits, color attributes, and texture attributes, are used to classify skin [9], [10]. However, classification based on inadequate feature sets can result in erroneous outcomes due to the high inter-class homogeneity and intra-class heterogeneity of skin lesions [11]. Conventional parametric methods require training data to be normally distributed, which is not the case for uncontrolled skin cancer data. Since every lesion has a different pattern, these traditional approaches are insufficient. As a result, deep learning methods have shown to be quite successful at classifying skin, helping dermatologists diagnose lesions with high precision. The application of deep learning in medicine has been widely explored through various detailed surveys. The most recent research on deep learning-based methods for melanoma detection and classification is covered in section 2.

Deep learning techniques based on transfer learning have gained widespread traction at the moment. With this method, large-scale datasets from one domain (like natural photographs) are used to train deep learning models, which are then used to transfer their learned representations to a target domain (biomedical images). By fine-tuning these pre-trained models on smaller biomedical datasets, researchers can effectively classify biomedical images with high accuracy. Transfer learning in biomedical image classification offers several advantages. Firstly, it allows researchers to overcome the challenge of limited annotated data in the biomedical domain by leveraging knowledge learned from larger datasets in related domains. Secondly, pre-trained models capture generic image features like textures, shapes, and edges, which can typically transferable across different image domains. This allows the model to learn important features for biomedical image classification tasks without requiring extensive retraining from scratch. Several models have been presented using this concept of transfer learning such as Zunair and Hamza [12] presented a model composed of two stages, combining adversarial training with transfer learning, Qureshi *et al.* [13] used Google Xception model to develop the transfer learning architecture, Hosny *et al.* [14] used Alexnet. However, the transfer learning-based models suffer from different issues such as domain transfer between source and target domain where resolution, noise level, and tissue variability affects the transferability of learned attributes. As a result, the pre-trained model is unable to capture the pertinent characteristics. Furthermore, a significant quantity of data is needed to refine the previously trained models. In this reserach, our main objective is to develop a deep learning-based melanoma categorization through the use of transfer learning strategy.

As discussed before, skin cancer occurs more commonly than any other form of cancer, with diagnosis rates surpassing all other types combined, and approximately 9,500 new cases are diagnosed each day in the United States alone. By 2040, skin cancer cases are projected to approach half a million, with melanoma being the deadliest type, marking a staggering 62% increase since 2018. Skin cancer is primarily caused by overexposure to UV radiation from sunlight or artificial sources like tanning machines, compounded by factors such as reduced ozone levels, geographical proximity to the equator, poor dietary habits, alcohol consumption, and smoking. Despite the urgency of the situation, current diagnostic methods like dermoscopy, biopsy, and macroscopic inspection are highly dependent on the clinician's skill, making accurate diagnosis challenging and time-consuming. Traditional computer image analysis algorithms and machine learning techniques have

attempted to assist in early diagnosis, but these methods often fall short due to the uncontrolled nature of skin cancer data and issues like inadequate feature sets, high inter-class homogeneity, and intra-class heterogeneity. Deep learning, especially transfer learning, offers a promising solution by leveraging large-scale datasets from related domains to enhance classification accuracy in biomedical images. However, transfer learning models face challenges such as domain transfer issues and the need for significant amounts of data to fine-tune pre-trained models effectively. This work focuses on developing a robust deep learning model for melanoma classification through transfer learning, aiming to overcome the limitations of existing approaches and improve the accuracy and reliability of early skin cancer detection.

The key contributions of this work are as follows: The pre-processing phase performs several tasks such as improving the image quality, hair removal and label encoding to process the efficiently. The proposed model uses transfer learning approach for feature extraction where VGG, InceptionResNet, ResNet and MobileNet modules are used to obtain the significant attributes. The obtained features are fused together to attain the final feature vector. In next step, deep ensemble model is constructed by using the concept of transfer learning where EfficientNet, Xception, and DenseNet transfer learning models are used to obtain the probability vector for predicting the skin cancer. Finally, FC and sigmoid layers are used to obtain the concluding classification.

This work introduces several key innovations that significantly enhance the performance of melanoma image classification systems. A comprehensive multi-stage pre-processing pipeline is implemented to address common dermatological imaging challenges. This pipeline improves contrast to highlight critical features and applies noise reduction to ensure cleaner and more reliable input for further analysis. To counteract dataset imbalance—a frequent issue in melanoma detection. Various data augmentation techniques are employed to synthetically expand the minority classes, enhancing the model's robustness and ability to generalize effectively. The approach also leverages pre-trained deep learning models like VGG, ResNet, MobileNet and Inception ResNet, for feature extraction. These models capture rich hierarchical representations, and by fusing their feature vectors, the system achieves a more comprehensive and discriminative understanding of the input images. Finally, ensemble classification using transfer learning is employed, combining the strengths of multiple architectures. This ensemble approach captures complex patterns and subtle distinctions within melanoma images, resulting in improved classification accuracy and overall system performance.

The subsequent sections of the article are structured as follows: section 2 provides a comprehensive literature review of current melanoma classification techniques. Section 3 presents an in-depth overview of the proposed model. Section 4 outlines the results of the proposed model along with a comparative analysis with current techniques. Lastly, section 5 offers closing thoughts and future directions for the research.

2. LITERATURE SURVEY

This section gives in detailed analysis of the current approaches for classifying melanoma with the help of deep learning and transfer learning methods. Lu and Zadeh [15] presented an automated method for utilizing dermoscopy images to diagnose skin cancer. To increase classification accuracy, the model used depth-wise separable convolution and the swish activation function, with XceptionNet serving as the base network. Jain *et al.* [16] presented a deep transfer learning model was introduced. To address the issue of data imbalance, the method included image data augmentation. Moreover, feature extraction and classification tasks were performed using transfer learning techniques such as VGG19, InceptionV3, and MobileNet, among others. Following these architectures, the task was completed by incorporating max pooling, flattening, a dense layer, and the sigmoid function.

Ali *et al.* [17] discussed the challenges associated with the current computerized skin lesion malignancy detection system due to various variables, such as uneven lesion sizes and shapes, different color illuminations, and varying light conditions. In the preprocessing stage, filters or kernels are employed to eliminate artifacts and noise, following which robust features are extracted using the feature extraction procedure. Finally, data augmentation is used to increase the size of the image collection and improve classification accuracy. The performance of this method is compared with that of other transfer learning models, including AlexNet and ReNet.

Balaha and Hassan [18] presented an automated melanoma classification and segmentation approach by using the sparrow search algorithm (SpaSA) and meta-heuristic optimization tool. The segmentation method involves employing 5 different U-Net models—U-Net, U-Net++, attention-based UNet, and several others—each with a unique configuration. Additionally, eight pre-trained convolutional neural network (CNN) models, including MobileNet with VGG (small, big), are utilized to optimize hyperparameters using the meta-heuristic SpaSA optimizer. Meswal *et al.* [19] introduced an ensemble strategy for classifying skin lesions using weighted averages. They obtained the weighted sum of transfer learning models including InceptionV3, VGG16, Xception, and ResNet50 to create the ensemble models.

Sadik *et al.* [20] proposed a CNN-based architecture that combines Xception and MobileNet. Specifically, they utilized CNN architectures commonly applied in computer vision applications, namely MobileNet and Xception, to develop a system to detect skin cancer. Karri *et al.* [21] introduced a novel strategy to address the issue with current deep learning techniques, namely the challenge of generalizing data from various sources, leading to domain shift even in well-trained deep learning models. To overcome this, they proposed a transfer learning methodology which involves two-phase cross-domain. This method utilizes the ImageNet and MoleMap datasets to construct and enhance data-level and model-level transfer learning models. Additionally, they introduced SknRSUNet for segmentation, which incorporates spatial attention features merging.

Shekar and Hailu [22] presented a deep transfer learning approach that combines six specially designed algorithms with the DenseNet-169 model to gather more detailed and richer features. Subsequently, the classification task is performed using the gradient boosting machine (GBM) classification model. Deng [23] highlighted the significance of challenging samples, suggesting that they contain crucial information. In their paper, they introduced a novel method called limited examples network (LSNet) aimed at recognizing and enhancing the learning of such difficult examples. LSNet utilizes a pseudo-inverse learning autoencoder with a patch-based structured input to quickly determine position-sensitive loss. By efficiently identifying position-sensitive loss, LSNet can recognize challenging samples effectively. Moreover, when dealing with skin lesion datasets with few samples, data augmentation is employed in conjunction with transfer learning to improve the accuracy of classification in deep learning models. Remya *et al*. [24] proposed a deep learning-based architecture integrating vision transformer, which combines channel attention and transfer learning techniques to deliver accurate region of interest (ROI) segmentation and classification.

3. PROPOSED MODEL

This section gives detail about the suggested deep transfer learning-based strategy to classify skin cancer. Several steps comprise the entire model, including feature extraction, augmentation, pre-processing, and classification. Figure 1 illustrates the general architecture of the proposed transfer learning-based methodology for categorizing skin cancer.

- Step 1: collection and loading the dataset. In this step, we consider the melanoma related publically available dataset for processing. The dataset consists of images, their corresponding ground truth for segmentation and labels for classes.
- Step 2: data pre-processing, normalization and label encoding. Generally, the skin cancer images have varied illumination which affects the image analysis tasks. Therefore, we apply contrast enhancement to obtain the refined dermoscopy image. Moreover, during capturing these images, the quality of images is degraded due to noise factor. To overcome this issue, we adopt image filtering modelto remove the noise. The melanoma affected data is imbalanced data therefore we also incorporate data augmentation methods to address the data imbalance issue. Finally, we apply label encoding mechanism to make it compatible with deep learning processing modules.
- Step 3: feature extraction and feature vector. Once the image data is pre-processed and labels are encoded appropriately, we perform feature extraction task by using pre-trained deep learning models. In this work, we have used VGG, ResNet, InceptionResNet, andMobileNetfor feature extraction by using their pre-trained weights. The obtained feature vectors of each models are fused together to formulate the final vector.
- Step 4: ensemble using transfer learning models for classification. In this step the resultant feature vector is processed through the training process which uses EfficientNet, Xception and DenseNet mocels are used to classify the image data.

3.1. Feature extraction

3.1.1. VGG

The VGG model is a deep CNN. The VGG architecture is constructed by using pooling layer, convolution layer, and fully connected layer. Consider I as the input image, and $VGG(I)$ as the features extracted from the VGG model. The convolution operation of this model can be articulated as shown in (1) and (2):

$$Z^{[l]} = Conv\left(A^{[l-1]}, W^{[l]}, b^{[l]}\right) \quad (1)$$

$$A^{[l]} = ReLU\left(Z^{[l]}\right) \quad (2)$$

Where l depicts the index of the layer, $A^{[l-1]}$ represents the activation from the previous layer, $W^{[l]}$ and $b^{[l]}$ are the weights and biases of l^{th} convolutional layer, respectively and $Z^{[l]}$ represents the output of convolution operation. The MaxPooling operation can be performed as shown in (3):

$$P^{[l]} = MaxPool\left(A^{(l)}\right) \tag{3}$$

Finally, in (4) to (6) shows the operation of the fully connected layer:

$$F = Flatten\left(P^{[l]}\right) \tag{4}$$

$$Z^{[l+1]} = F.W^{[l+1]} + b^{[l+1]} \tag{5}$$

$$VGG(I) = ReLU(Z^{[l+1]} \tag{6}$$

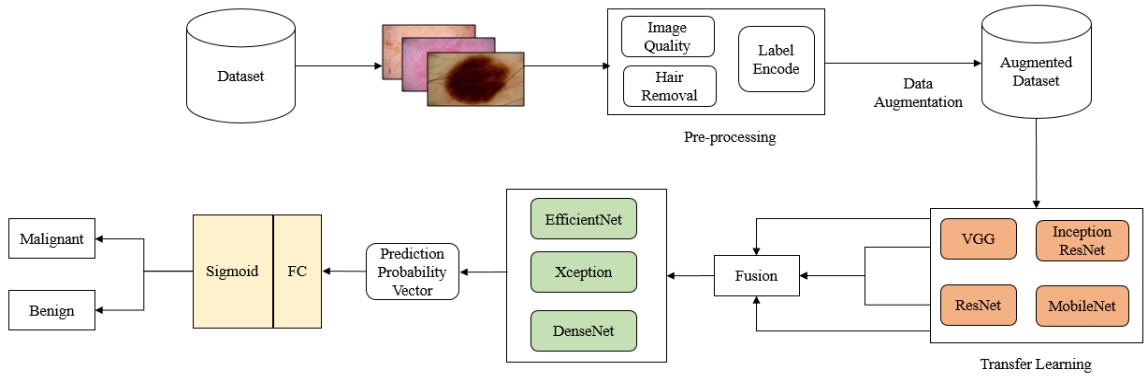


Figure 1. Architecture of proposed model

3.1.2. ResNet model

A deep CNN architecture called residual network (ResNet) was created to solve the issue of vanishing gradients. Instead of teaching the network the underlying mappings directly, it introduces skip connections, which enable the network to learn residual functions. This enables training of much deeper networks with improved performance. In order to perform the feature extraction task for melanoma images, let us consider that input image is denoted as I and the features extracted from the ResNet model as $ResNet(I)$. For feature extraction, this model uses convolution, residual block, pooling layer, and fully connected layer. The (7) and (8) express ResNet's convolution layer action as:

$$Z^{[l]} = Conv\left(A^{[l-1]}, W^{[l]}, b^{[l]}\right) \tag{7}$$

$$A^{[l]} = BatchNorm\left(ReLu\left(Z^{[l]}\right)\right) \tag{8}$$

Where l represents layer's index, $A^{[l-1]}$ represents the activation from the previous layer, $W^{[l]}$ denotes l^{th} weight convolution layer, $b^{[l]}$ denotes biases of l^{th} convolutional layer, and $Z^{[l]}$ depicts output of convolution operation. In next phase, it performs residual block operation which can be written as (9):

$$A^{[l]} = A^{[l-1]} + F\left(A^{[l-1]}, W^{[l]}\right) \tag{9}$$

Where F represents the residual function to be learned by the residual block, $A^{[l-1]}$ represents the input activation to the residual block. The addition operation denotes the skip connection, allowing the network to learn residual functions. Later, pooling operation is performed as (10):

$$P^{[l]} = MaxPool\left(A^{[l]}\right) \tag{10}$$

Where $P^{[l]}$ is the outcome of max pooling operation. Finally, the fully connected layer operations are performed as (11) to (13):

$$F = Flatten\left(P^{[l]}\right) \tag{11}$$

$$Z^{[l+1]} = F.W^{[l+1]} + b^{[l+1]} \tag{12}$$

$$VGG(I) = SoftMax\left(Z^{[l+1]}\right) \tag{13}$$

3.1.3. Inception ResNet

The Inception ResNet architecture combines Inception modules and residual connections from ResNet, resulting accurate and efficient CNN. It intends to use the advantages of both architectures to achieve improved performance on various computer vision tasks. This model consists of inception module, residual connection, pooling layer, and fully connected layer. The Inception module has various parallel convolutional branches with different kernel sizes and pooling operations. Each branch captures features at different scales and resolutions. It can be represented as (14):

$$Inception\left(A^{[l-1]}\right) = [Branch_1, Branch_2, \dots, Branch_n] \tag{14}$$

Where $A^{[l-1]}$ is activation from previous layer. Similar to ResNet, InceptionResNet incorporates residual connections within its architecture to facilitate training of very deep networks. This operation can be expressed as (15):

$$A^l = A^{[l-1]} + F\left(A^{[l-1]}\right) \tag{15}$$

Where F represents the residual function to be learned by the residual connection. In next step, we apply pooling operation similar to ResNet and VGG model as shown in (16):

$$P^{[l]} = MaxPool\left(A^{[l]}\right) \tag{16}$$

Finally, it uses fully connected layer operations as discussed in VGG and ResNet models which is expressed as (17) to (19):

$$F = Flatten\left(P^{[l]}\right) \tag{17}$$

$$Z^{[l+1]} = F.W^{[l+1]} + b^{[l+1]} \tag{18}$$

$$VGG(I) = SoftMax\left(Z^{[l+1]}\right) \tag{19}$$

3.1.4. MobileNet

MobileNet is a lightweight CNN architecture that utilizes depth wise separable convolutions. It helps to minimize the required number of parameters for training resulting in reduced computation time and maintaining high accuracy. Let's denote the input image as I, and the features extracted from the MobileNet model as $MobileNet(I)$. The feature extraction process includes depth wise and point wise convolution operations. The depth wise convolution can be articulated as (20) and (21):

$$Z^{[l]} = DepthwiseConv\left(A^{[l-1]}, W^{[l]}, b^{[l]}\right) \tag{20}$$

$$A^{[l]} = BatchNorm\left(ReLU\left(Z^{[l]}\right)\right) \tag{21}$$

Where l represents layer's index, $A^{[l-1]}$ represents the activation from the previous layer, $W^{[l]}$ denotes l^{th} weight convolution layer, $b^{[l]}$ denotes biases of l^{th} convolutional layer, and $Z^{[l]}$ represents the output of depthwise convolution operation. The next step performs, Pointwise convolution operation which is expressed as (22) and (23):

$$Z^{[l+1]} = PointwiseConv\left(A^{[l]}, W^{[l+1]}, b^{[l+1]}\right) \tag{22}$$

$$MobileNet(I) = Softmax\left(Z^{[l+1]}\right) \tag{23}$$

The final feature vector can be obtained by concatenating these attributes and normalizing the concatenated feature vector. It can be expressed as (24):

$$F = normalize[Concatenate(VGG(I), ResNet(I), InceptionResNet(I), MobileNet(I) \quad (24)$$

3.2. Ensemble deep transfer learning model

This section presents the ensemble model of deep transfer learning models for melanoma classification. The proposed model uses stacked ensemble of EfficientNetB0, Xception, and DenseNet121 to obtain he classification probability vector. Finally, the fully connected and sigmoid layers are used to obtain the final classification based on the concatenation of initial probability vector. In Figure 2 depicts the stacked ensemble classification model.

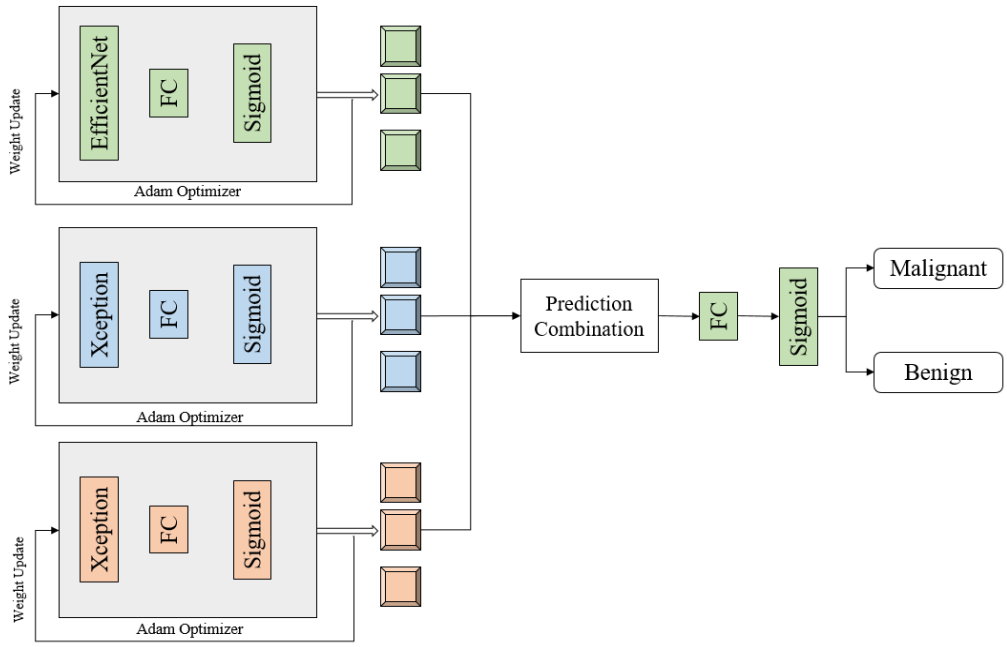


Figure 2. Proposed deep transfer learning module for melanoma classification

Efficient Net is a family of CNN architectures that have been designed to achieve the improved performance with less computational parameters compared to other architectures. In this work, we have used EfficientNetB0 which is the baseline model in the EfficientNet family. It performs convolution operations, depth wise separable convolution, global average pooling and fully connected operation. The feature extraction model can be articulated as (25) and (26):

$$Z^{[l]} = Conv\left(A^{[l-1]}, W^{[l]}, b^{[l]}\right) \quad (25)$$

$$A^{[l]} = Swish\left(BatchNorm\left(Z^{[l]}\right)\right) \quad (26)$$

Where l represents layer's index, $A^{[l-1]}$ represents the activation after applying batch normalization and the swish activation function, $W^{[l]}$ is weights, $b^{[l]}$ is biases of l^{th} convolutional layer, and $Z^{[l]}$ denotes the output of convolution operation. Next step, it performs global average pooling (GAP) as in (27):

$$G = GAP\left(A^{[l]}\right) \quad (27)$$

Finally, the fully connected layer is applied as (28) and (29):

$$Z^{]l+1]} = G.W^{[l+1]} + b^{[l+1]} \quad (28)$$

$$EfficientNetB0(I) = Softmax\left(Z^{[l+1]}\right) \quad (29)$$

Here, Softmax denotes the softmax activation function, which converts the raw scores into class probabilities. Similarly, Xception and Densenet models also used to construct the classifier where Softmax function converts the raw score into class probabilities.

4. RESULTS AND DISCUSSION

A detailed study of the recommended approach is given in this section. The dataset utilised in this work is described in depth in the first subsection. The next subsection explains the metrics that were used to assess the proposed work's performance. Finally, a comparison analysis of proposed work with existing techniques is done.

4.1. Dataset details

Research on the classification of melanoma has advanced significantly due to the attention from International Skin Imaging Collaboration (ISIC) challenges. High resolution skin lesion picture collections with expert annotation, biopsy-proven, and global information are made available by the ISIC. The organisation has conducted annual skin lesion challenges in an effort to boost researcher engagement and improve CAD algorithm detection. Table 1 provides an overview of the ISIC datasets from 2016 to 2020.

- ISIC 2016 dataset: there are 1279 photos in total in this dataset, 900 samples are used for training and 379 are used for testing. Both the training and testing sets have access to the ground truth data.
- ISIC 2017 dataset: in this dataset 2600 images are available of which 2000 are usesd for network testing and 600 for training. This dataset contains the groundtruths for four different class groups: melanoma, seborrhoeic keratosis, nevus, and melanoma nevus
- ISIC 2018 dataset: this large dataset has a total of 11,527 photos, of which 10,015 were used for network training and the remaining 1512 for network testing and performance evaluation.
- ISIC 2019 dataset: of the 33,569 photos in the ISIC 2019 dataset, there are 8,238 images in the testing set and 25,331 images in the training set. But just the labels for the training set photos, which represent eight classes, are included in this dataset. The information about the training and testing photos is contained in the metadata. The training metadata contains all pertinent patient information, whereas the testing metadata contains the age, gender, anatomical site, and lesion ID of the patient.
- ISIC 2020 dataset: there are 44108 photos in total in this dataset; 33126 are used for training and 10982 are used for testing. The ground truth data, which includes lesion ID, gender, age, patient ID, diagnosis, anatomical place, and benign or malignant status, is supplied for the training set, just like it was the year before.
- ISIC Kaggle: in this work, Kaggle dataset of skin lesions which is publicly available from the ISIC library, is used to train and validate our stacking ensemble model. There are 1800 benign and 1497 malignant mole photos in the collection [14]. Upon closer inspection, sounds and artefacts were discovered in the gathered skin lesion photos. We used, standard pre-processing methods like scaling, normalisation, noise reduction, and contrast enhancement, to address this.

The pixel intensity value range of evey image is normalized set to [0,1]. The images are then uniformly enlarged to 224 by 224 dimensions. The sample image of Benign and Malignant melanoma cases is shown in Figures 3(a) and 3(b) respectively, these images are obtained from ISIC challenge datasets.

Table 1. Summary of ISIC 2016-2020 dataset

Dataset	Train	Test	Total
ISIC 2016	900	379	1279
ISIC 2017	2000	600	2600
ISIC 2018	10015	1512	11527
ISIC 2019	25331	8238	33569
ISIC 2020	33126	10982	44108

4.2. Performance measurement parameters

This subsection describes the performance measurement parameters that were used to analyze the overall performance of the proposed model. The classifier's performance is evaluated using the confusion matrix, which displays the quantity of true negatives (TN), and false negatives (FN), true positives (TP) and false positives (FP). Figure 4 provides a basic illustration of a confusion matrix.

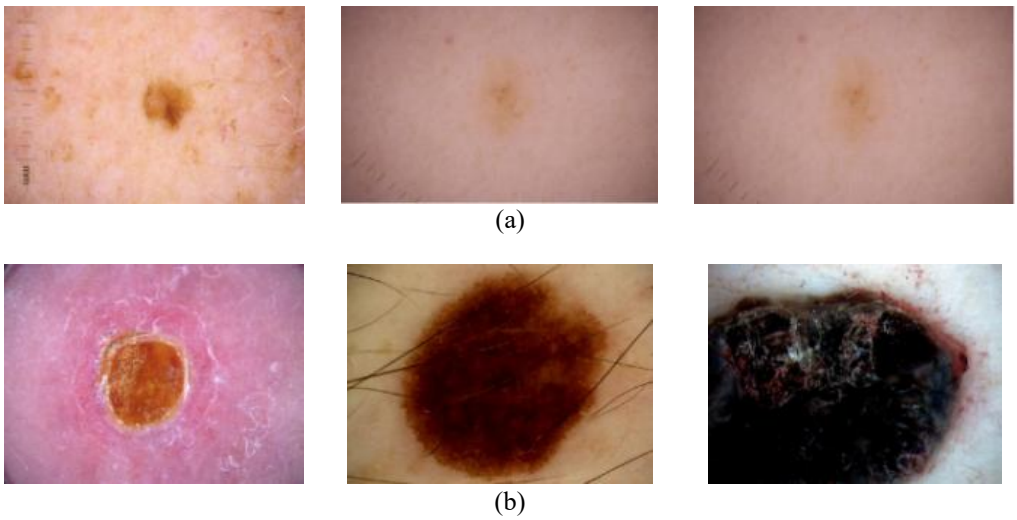


Figure 3. The sample image of melanoma cases for (a) benign samples (b) malignant samples

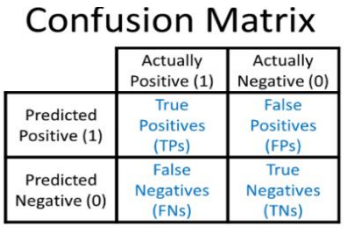


Figure 4. Representation of confusion matrix

To evaluate the effectiveness of our proposed method, we employ established metrics including accuracy, F-measure, precision, false positive rate, true positive rate. These measures depend on true positive, true negative, false positive, and false negative being distinguished. In the context of this study, TP, TN, FP, and FN represent properly recognized malignant images, correctly classed benign images, wrongly categorized benign images, and incorrectly classified malignant images, respectively.

The ratio of TP to the total number of pictures classified as malignant is used to calculate precision is shown in (30).

$$Precision = \frac{TP}{TP+FP} \tag{30}$$

Divide the total number of harmful pictures by the true positive rate (TPR) as in (31):

$$TPR = \frac{TP}{TP+FN} \tag{31}$$

The false positive rate (FPR) is calculated as in (32):

$$FPR = \frac{FP}{FP+TN} \tag{32}$$

The definition of accuracy is the product of TN and TP divided by overall pictures, as written in (33):

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{33}$$

F-measure in (34) represents the harmonic mean of precision and recall:

$$F-measure = 2 \times \frac{Precision \times Recall}{(Precision+Recall)} \tag{34}$$

4.3. Comparative performance analysis

This subsection shows result of the proposed model along with a performance comparison between the obtained results and current systems. Table 2 displays the comparison analysis using several deep learning-based techniques and datasets. Table 2 shows hows well the proposed model performs alongside cutting-edge melanoma classifiers. A deep learning model was presented in [25] with a classification accuracy of 0.92. A new VGG-13 model for skin cancer diagnosis was given by Gilani *et al.* [26], and it achieved an 89.57% detection accuracy. Based on Inception-V3, the ConvNet model introduced in [27] focuses on binary classification of skin conditions and successfully differentiates between benign and malignant types of skin cancer. Malik *et al.* [28] showcased the multi-classification of skin lesions using 2D superpixels with ResNet-50, achieving an accuracy of 85.50%. Ling *et al.* [29] achieved a precision of 88.10% in the multi-classification of skin cancer. Zhou *et al.* [30] presented SCDNet and achived accuracy of 92.89% in classification of skin cancer. In contrast to established well known methods, our proposed model demonstrates an improved accuracy for ISIC 2016-2020 datasets.

Table 2. Dataset comparative analysis with different datasets

Article	Year	Model	Dataset	Accuracy	Recall	Precision	F1-score
[25]	2023	CNN	ISIC-2017	92.01	91.91	91.66	91.99
[26]	2023	VGG-13	ISIS-2019	89.58	90.69	89.65	89.63
[27]	2021	ConvNet	ISIC-2018	86.89	86.15	87.50	-
[28]	2022	RCNN + 2Dsuperpixel	HAM-10000	85.49	83.39	84.49	85.30
[29]	2021	ResetXt101	ISIC-2019	88.49	87.39	88.10	88.30
[30]	2022	SCDNet	ISIC-2019	92.89	92.20	92.19	92.20
		Proposed	ISIC 2016	96.10	95.15	96.25	95.20
			ISIC 2017	97.23	96.20	96.30	95.55
			ISIC 2018	97.50	97.88	97.50	98.20
			ISIC 2019	98.33	98.50	98.30	98.15
			ISIC 2020	98.60	98.90	98.50	97.30

In the following experiment, we compared the accuracy of the proposed approach to the most advanced deep learning, machine learning, and transfer learning techniques. The Table 3, shows the comparative analysis for HAM-10000. As discussed before, the transfer learning models have gained huge attention in this biomedical imaging domain therefore several transfer learning-based models have been introduced. To assess the efficiency of these transfer learning models, we evaluated the performance for Kaggle ISIC dataset. Table 4 demonstrates the outcome of mostly used transfer learning models for image classification tasks.

Table 3. Comparative analysis for HAM10000 dataset

Article	Year	Model	Dataset	Accuracy (%)
[31]	2020	AlexNet	HAM-10000	84
[32]	2019	MobileNet	HAM-10000	83.9
[33]	2020	MobileNet, VGG-16	HAM-10000	80.61
[34]	2019	SVM	HAM-10000	74.75
[35]	2020	ResNet	HAM-10000	78
	2020	Xception		82
	2020	DenseNet		82
[36]	2020	CNN	HAM-10000	77
[37]	2021	MobileNet and LSTM	HAM-10000	85
[38]	2021	Inception-V3	HAM-10000	89.73
[39]	2023	InceptionResnet-V2	HAM-10000	91.26
		Transfer Learning	HAM-10000	98.55

Table 4. Overall performance analysis for Kaggle ISIC dataset

Model	Accuracyc	Precision	Recall	Specificity	F1-score	AUC
ResNet50	88.78	93.33	85.56	92.67	89.28	0.891
VGG-16	90.91	95.68	86.11	95.33	90.64	0.907
Xception	92.42	93.30	92.78	92.00	93.04	0.924
DenseNet 121	92.27	91.87	94.17	90.00	93.00	0.921
EfficientNetBO	92.30	94.02	91.67	93.00	92.83	0.923
Proposed Model	98.76	98.60	97.67	95.67	97.13	0.987

This experiment shows that, proposed method reports, 95.76% of overall accuracy and average precision is recorded at 98.60%. Figure 5, shows a graphic representation of the attained performance characteristics in terms of precision, accuracy, specificity, F1-score, recall, and area under curve. For this experiment, the training loss and validation loss performance is shown in Figure 6. The Figure 6(a) shows the performance of validation loss and training loss, and it is observed that gretest performance is reported at 9 epochs. Similarly, Figure 6(b) shows the performance of validation accuracy and training accuracy and it is observed that highest accuracy is reported at 15th epoch. This thorough research shows that the suggested model not only outperforms several cutting-edge transfer learning techniques but also conventional deep learning models, setting a new standard for melanoma classification accuracy and overall performance.

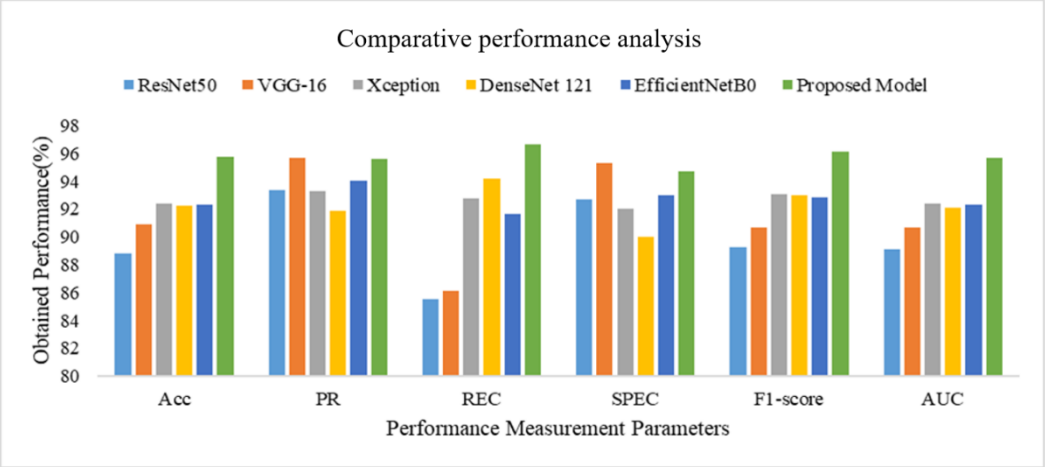
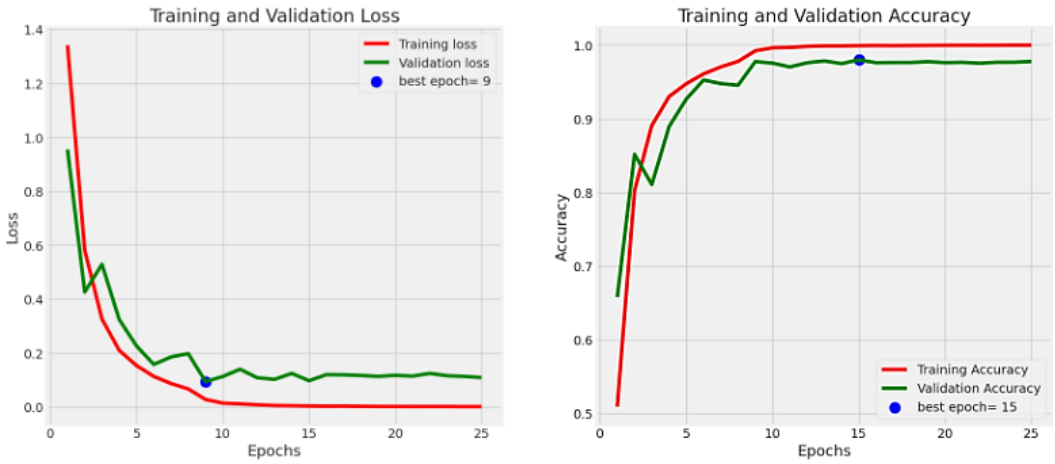


Figure 5. Comparative performance for varied transfer learning models



(a) (b)

Figure 6. Comparative performance of training and validation for (a) loss and (b) accuracy

5. CONCLUSION

In this work, we address the pressing need for improved early detection and diagnosis of melanoma, a formidable challenge in public health due to its increasing mortality rates. Machine learning has become a viable way to improve accuracy, even though more conventional techniques, such as visual inspection, mainly depend on observer expertise. However, researchers continue to face difficulties in reaching high accuracy. To overcome this challenge, we proposed a comprehensive approach utilizing deep transfer learning techniques. Our methodology encompasses several crucial steps to enhance the processing of melanoma-related image datasets. Through data pre-processing, normalization, and label encoding, we mitigate issues such as varied illumination, image noise, and data imbalance. Furthermore, we leverage pre-

trained deep learning models like ResNet, MobileNet, VGG, and InceptionResNet for feature extraction, followed by ensemble learning using transfer learning models like EfficientNet, Xception, and DenseNet for classification. By fusing feature vectors and employing ensemble learning, our goal is develop a reliable accurate model for melanoma classification. Our proposed approach demonstrates promising results on publicly available ISIC datasets from 2017 to 2020. The reported average accuracies of 96.10%, 97.23%, 97.50%, 98.33%, and 98.60%, respectively, underscore the effectiveness of our methodology in accurately identifying melanoma. These findings represent a substantial development in the profession and could result in better patient outcomes and earlier detection. In future work, this work can be extended to provide a universal architecture to identify differnet types of diseases.

FUNDING INFORMATION

This research received no external funding.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Soumya Gadag	✓	✓	✓	✓	✓	✓		✓	✓	✓				
Panduranga Rao Malode Vishwanatha		✓			✓	✓	✓	✓		✓		✓	✓	
Virupaxi Balachandra Dalal	✓		✓	✓	✓	✓				✓	✓			

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY

The data are publicly available in the International Skin Imaging Collaboration (ISIC) at https://www.isic-archive.com/ and Kaggle at https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000.

REFERENCES

[1] Skin Cancer Foundation, "Skin cancer facts and statistics," *Skin Cancer Foundation*. 2025. Accessed: Apr. 18, 2024. [Online]. Available: https://www.skincancer.org/skin-cancer-information/skin-cancer-facts
[2] Melanoma UK, "2020 melanoma skin cancer report," *Melanoma UK*. 2020. Accessed: Apr. 18, 2024. [Online]. Available: https://www.melanomauk.org.uk/2020-melanoma-skin-cancer-report
[3] E. J. Orrin, P. B. Cassidy, R. P. Kulkarni, E. G. Berry, and S. A. Leachman, "Melanoma prevention," in *Melanoma in Clinical Practice*, Cham, Switzerland: Springer, 2021, pp. 3–29, doi: 10.1007/978-3-030-82639-0_1.
[4] NHS, "Causes: Melanoma skin cancer" *National Health Services.* 2023. Accessed: Apr. 20, 2024. [Online]. Available: https://www.nhs.uk/conditions/melanoma-skin-cancer/causes/
[5] WHO, "More can be done to restrict sunbeds to prevent increasing rates of skin cancer," *World Health Organization*. 2017. Accessed: Apr. 20, 2024. [Online]. Available: https://www.who.int/news/item/21-06-2017-more-can-be-done-to-restrict-sunbeds-to-prevent-increasing-rates-of-skin-cancer
[6] Department for Environment Food Rural Affairs, "Depletion of the ozone layer leading to an increase in ground-level ultraviolet radiation," *UK Air Information Resource.* Accessed: Apr. 20, 2024. [Online]. Available: https://uk-air.defra.gov.uk/research/ozone-uv/moreinfo?view=increase-uv-radiation
[7] R. Kaur, H. GholamHosseini, R. Sinha, and M. Lindén, "Melanoma classification using a novel deep convolutional neural network with dermoscopic images," *Sensors*, vol. 22, no. 3, 2022, doi: 10.3390/s22031134.
[8] A. Masood and A. A. Al-Jumaily, "Computer aided diagnostic support system for skin cancer: A review," *International Journal of Biomedical Imaging*, vol. 2013, 2013, doi: 10.1155/2013/323268.

[9] A. S. Alphonse and M. S. Starvin, "A novel and efficient approach for the classification of skin melanoma," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 12, pp. 10435–10459, 2021, doi: 10.1007/s12652-020-02648-x.
[10] R. D. Seeja and A. Suresh, "Melanoma classification employing inter neighbor statistical color and mean order pattern texture feature," *Multimedia Tools and Applications*, vol. 80, no. 13, pp. 20045–20064, 2021, doi: 10.1007/s11042-021-10685-7.
[11] I. Iqbal, M. Younus, K. Walayat, M. U. Kakar, and J. Ma, "Automated multi-class classification of skin lesions through deep convolutional neural network with dermoscopic images," *Computerized Medical Imaging and Graphics*, vol. 88, 2021, doi: 10.1016/j.compmedimag.2020.101843.
[12] H. Zunair and A. B. Hamza, "Melanoma detection using adversarial training and transfer learning," *Physics in Medicine & Biology*, vol. 65, no. 13, 2020, doi: 10.1088/1361-6560/ab86d3.
[13] M. N. Qureshi, M. S. Umar, and S. Shahab, "A transfer-learning-based novel convolution neural network for melanoma classification," *Computers*, vol. 11, no. 5, 2022, doi: 10.3390/computers11050064.
[14] K. M. Hosny, M. A. Kassem, and M. M. Fouad, "Classification of skin lesions into seven classes using transfer learning with AlexNet," *Journal of Digital Imaging*, vol. 33, pp. 1325–1334, 2020, doi: 10.1007/s10278-020-00371-9.
[15] X. Lu and Y. F. A. Zadeh, "Deep learning-based classification for melanoma detection using XceptionNet," *Journal of Healthcare Engineering*, vol. 2022, 2022, doi: 10.1155/2022/2196096.
[16] S. Jain *et al.*, "Deep learning-based transfer learning for skin cancer classification," *Sensors*, vol. 21, no. 23, 2021, doi: 10.3390/s21238142.
[17] M. S. Ali, M. S. Miah, J. Haque, M. M. Rahman, and M. K. Islam, "An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models," *Machine Learning with Applications*, vol. 5, 2021, doi: 10.1016/j.mlwa.2021.100036.
[18] H. M. Balaha and A. E. S. Hassan, "Skin cancer diagnosis based on deep transfer learning and sparrow search algorithm," *Neural Computing and Applications*, vol. 35, no. 1, pp. 815–853, 2023, doi: 10.1007/s00521-022-07762-9.
[19] H. Meswal *et al.*, "A weighted ensemble transfer learning approach for melanoma classification from skin lesion images," *Multimedia Tools and Applications*, pp. 1–23, 2023, doi: 10.1007/s11042-023-16783-y.
[20] R. Sadik, A. Majumder, A. A. Biswas, B. Ahammad, and Md. M. Rahman, "An in-depth analysis of convolutional neural network architectures with transfer learning for skin disease diagnosis," *Healthcare Analytics*, vol. 3, 2023, doi: 10.1016/j.health.2023.100143.
[21] M. Karri, C. S. R. Annavarapu, and U. R. Acharya, "Skin lesion segmentation using two-phase cross-domain transfer learning framework," *Computer Methods and Programs in Biomedicine*, vol. 231, 2023, doi: 10.1016/j.cmpb.2023.107408.
[22] B. H. Shekar and H. Hailu, "Fusion of features extracted from transfer learning and handcrafted methods to enhance skin cancer classification performance," in *CVMI 2022*, Springer, 2023, pp. 243–257, doi: 10.1007/978-981-19-7867-8_20.
[23] X. Deng, "LSNet: a deep learning based method for skin lesion classification using limited samples and transfer learning," *Multimedia Tools and Applications*, vol. 83, pp. 61469–61489, 2024, doi: 10.1007/s11042-023-17975-2.
[24] S. Remya, T. Anjali, and V. Sugumaran, "A novel transfer learning framework for multimodal skin lesion analysis," *IEEE Access*, vol. 12, pp. 50738-50754, 2024, doi: 10.1109/access.2024.3385340.
[25] S. Wang, Y. Xing, L. Zhang, H. Gao, and H. Zhang, "Deep convolutional neural network for ulcer recognition in wireless capsule endoscopy: experimental feasibility and optimization," *Computational and Mathematical Methods in Medicine*, vol. 2019, no. 1, 2019, doi: 10.1155/2019/7546215.
[26] S. Q. Gilani, T. Syed, M. Umair, and O. Marques, "Skin cancer classification using deep spiking neural network," *Journal of Digital Imaging*, vol. 36, no. 3, pp. 1137-1147, doi: 10.1007/s10278-023-00776-2.
[27] F. Wen and A. K. David, "A genetic algorithm based method for bidding strategy coordination in energy and spinning reserve markets," *Artificial Intelligence in Engineering*, vol. 15, pp. 71–79, 2001, doi: 10.1016/S0954-1810(01)00002-4.
[28] H. Malik, M. S. Farooq, A. Khelifi, A. Abid, J. N. Qureshi, and M. Hussain, "A comparison of transfer learning performance versus health experts in disease diagnosis from medical imaging," in *IEEE Access*, vol. 8, pp. 139367-139386, 2020, doi: 10.1109/ACCESS.2020.3004766.
[29] W. Ling, X. Wang, J. Fu, and L. Zhen, "A novel probability binary particle swarm optimization algorithm and its application," *Journal of Software*, vol. 3, no. 9, pp. 28–35, 2008, doi: 10.4304/jsw.3.1.28-35.
[30] Y. Zhou, C. Koyuncu, C. Lu, R. Grobholz, I. Katz, A. Madabhushi, and A Janowczyk, "Multi-site cross-organ calibrated deep learning (MuSClD): Automated diagnosis of non-melanoma skin cancer," *Medical Image Analysis*, vol. 84, 2023, doi: 10.1016/j.media.2023.
[31] A. Ameri, "A deep learning approach to skin cancer detection in dermoscopy images," *Journal of Biomedical Physics and Engineering*, vol. 10, pp. 801–806, 2020, doi: 10.31661/jbpe.v0i0.2004-1107.
[32] W. S. -Lim, W. Wettayaprasit, and P. Aiyarak, "Convolutional neural networks using MobileNet for skin lesion classification," in *2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, 2019, pp. 242-247, doi: 10.1109/JCSSE.2019.8864113.
[33] A. C. Salian, S. Vaze, P. Singh, G. N. Shaikh, S. Chapaneri, and D. Jayaswal, "Skin lesion classification using deep learning architectures," *2020 3rd International Conference on Communication System, Computing and IT Applications (CSCITA)*, Mumbai, India, 2020, pp. 168-173, doi: 10.1109/CSCITA47329.2020.9137810.
[34] T. C. Pham, G. S. Tran, T. P. Nghiem, A. Doucet, C. M. Luong, and V. -D. Hoang, "A comparative study for classification of skin cancer," *2019 International Conference on System Science and Engineering (ICSSE)*, Dong Hoi, Vietnam, 2019, pp. 267-272, doi: 10.1109/ICSSE.2019.8823124.
[35] Z. Rahman and A. M. Ami, "A transfer learning based approach for skin lesion classification from imbalanced data," *2020 11th International Conference on Electrical and Computer Engineering (ICECE)*, Dhaka, Bangladesh, 2020, pp. 65-68, doi: 10.1109/ICECE51571.2020.9393155.
[36] K. Polat and K. O. Koc, "Detection of skin diseases from dermoscopy image using the combination of convolutional neural network and one-versus-all," *Journal of Artificial Intelligence Systems*, vol. 2, pp. 80–97, 2020, doi: 10.33969/JAIS.2020.21005.
[37] P. N. Srinivasu, J. G. SivaSai, M. F. Ijaz, A. K. Bhoi, W. Kim, and J. J. Kang, "Classification of skin disease using deep learning neural networks with MobileNet V2 and LSTM," *Sensors*, vol. 21, no. 8, 2021, doi: 10.3390/s21082852.
[38] N. Kausar *et al.,* "Multiclass skin cancer classification using ensemble of fine-tuned deep learning models," *Applied Sciences*, vol. 11, 2021, doi: 10.3390/app112210593.
[39] G. Alwakid, W. Gouda, M. Humayun, and N. Z. Jhanjhi, "Diagnosing melanomas in dermoscopy images using deep learning," *Diagnostics*, vol. 13, no. 10, 2023, doi: 10.3390/diagnostics13101815.

BIOGRAPHIES OF AUTHORS

Soumya Gadag completed her B.E. degree in Electronics and Communication Engineering from Visvesvaraya Technological University, India in 2008. She received her M.Tech. from Visvesvaraya Technological University in 2012. Currently working as Associate Professor in the Department of Electronics and Communication Engineering, Jain College of Engineering and Research, India. Her research interests include wireless communication, image processing, machine learning, and artificial intelligence. She can be contacted at email: soumya.gadag5@gmail.com.

Panduranga Rao Malode Vishwanatha obtained his Ph.D. from the National Institute of Technology Karnataka, Mangalore, India. He has completed a Master of Technology in Computer Science and a Bachelor of Engineering in Electronics and Communication. He works as a Professor in Jain (Deemed to be University) Bengaluru, India. His research interests are in the field of real-time and embedded systems on Linux platforms. He has published various research papers in journals and conferences across India. He visited JAPAN in 2008 for the IEEE international conference in Okinawa Island. He has authored two reference books on Linux internals. He is a life member of the Indian Society for Technical Education and IAENG. In the past three years, he has published 12 Indian patents. Here, three patents are stepping towards award/grant status. He can be contacted at email: r.panduranga@jainuniversity.ac.in.

Virupaxi Balachandra Dalal is currently working as Professor in the Department of Electronics and Communication Engineering at Jain College of Engineering and Research, Belagavi, Karnataka, India. He received his doctoral degree from Visveswaraya Technological University, Belagavi, Karnataka, India. His area of interest is signals processing and image processing, wireless communication, and digital electronics. He can be contacted at email: virupaxidalal@gmail.com.