International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Deep Learning Enhanced Skin Cancer Analysis: Advancements in Melanoma Detection with Medical Support

Dr. Madheswari K [1], Dr. Karthikeyan N [2], Janardhan Kantubhukta[3]

Submitted: 29/01/2024 **Revised**: 07/03/2024 **Accepted**: 15/03/2024

Abstract: The identification and classification of skin cancer lesions are crucial due to their profound health implications. The paper presents web application based a deep learning model to address the skin cancer detection. This proposed system aimed to fine-tune well-established models for skin lesion classification. Ensemble learning techniques were applied by combining thoroughly fine-tuned Inception V3 and DenseNet201 models to improve classification accuracy. In this web based model development, the front end model is developed to provide user interface for skin cancer classification using Python, JavaScript, HTML, CSS and Flask framework. The proposed scheme is compared to other standard models such as VGG-16, Inception V3, Inception ResNet V2, and DenseNet201. In the experimental part, the proposed scheme achieves higher accuracy rate of 88.5% than compared to Baseline CNN model with 75.64 %, VGG-16 with 79.65 %, Inception ResNet V2 with 82.50 %, with HAM10000 dataset, showcasing the efficacy of our methodology in precisely categorizing skin cancer lesions.

Keywords: *Skin cancer, Deep learning, Fine-tuning, VGG-16, Inception V3, Inception ResNet V2, DenseNet201, Ensemble learning, Web based model.*

1. Introduction

Skin cancer presents a significant public health concern, impacting individuals globally. Its prevalence has been consistently rising, warranting significant attention in public health discussions.[1] Basal cell carcinoma, melanoma, and squamous cell carcinoma emerge as the most frequently diagnosed forms of skin cancer, each posing distinct challenges in terms of diagnosis and treatment.

Prompt identification and precise categorization of skin lesions play a pivotal role in facilitating timely interventions and improving patient prognosis. Traditionally, dermatologists have relied on visual inspection and histopathological analysis for diagnosing skin cancer. Yet, this method may involve subjectivity, consume considerable time, and be susceptible to errors, emphasizing the necessity for objective and efficient diagnostic methods.

Recent progress in deep learning and computer vision has ushered in a new era of automated systems, enabling the precise and swift analysis of medical images, such as skin lesions, with remarkable accuracy. Convolutional neural networks (CNNs) have developed as particularly effective tools for image classification tasks, leveraging their ability to extract intricate patterns and structures from images through hierarchical feature analysis. Our proposed work leverage deep learning techniques for the accurate classification of skin lesions using the HAM10000 dataset comprises an extensive assortment of dermatoscopic images encompassing diverse pigmented skin lesions. One of the key challenges in analysing dermatoscopic images is the presence of artifacts such as hair, which can obscure the underlying lesion and hinder accurate classification. Our study implement a pre-processing step to address this challenge to remove hair from the images. By mitigating the impact of artifacts such as hair, our pre-processing approach aims to improve the quality of input data for subsequent classification tasks. Our proposed work hypothesize that cleaner and more refined images will enable our deep-learning models to extract relevant features more accurately.

In this context, our research endeavours to leverage the capabilities of deep learning and Convolutional neural networks (CNNs), to solve the challenges associated with skin cancer detection and classification. By fine-tuning pre-trained CNN models, such as VGG-16, Inception V3, Inception ResNet V2, and DenseNet201, Our proposed work aim to develop a reliable model for automatic identification and categorizing skin lesions.

Furthermore, our study explore the potential of ensemble learning techniques to enhance the performance of our classification system by combining multiple models and leveraging their complementary strengths. Additionally, our study develop a user-friendly web application to

[1] *Associate Professor, School Of Computer science and Engineering, Vellore institute of technology, Chennai-600127, India.*
Email: madheswari.k@vit.ac.in
[2] *Associate Professor, School Of Computer science and Engineering, Vellore institute of technology, Chennai-600127, India.*
Email: narenkarthikeyan.mecse@gmail.com
[3] *School of Computer science and Engineering, Vellore institute of technology, Chennai-600127, India.*
Email: kantubhukta.janardhan2020@vitstudent.ac.in

democratize access to skin cancer classification, providing a seamless interface for healthcare professionals and individuals alike.

2. Literature Survey

Doaa Khalid Abdulridha Al-Saedi and Serkan Savaş [2] Compared skin cancer classification using DenseNet, Xception, InceptionResNetV2, ResNet50, and MobileNet on dermatoscopic image datasets, notably the ISIC dataset. Augmenting photos addressed dataset imbalances, resulting in 98.35% accuracy with modified DenseNet121. Challenges include dataset imbalances and variable model performance based on dataset size. Despite hurdles, deep transfer learning networks hold promise in skin cancer classification, prompting ongoing research to address dataset imbalances and optimize model performance.

Karar Ali et al [3] done skin cancer classification utilizing EfficientNets surpassed dermatologists visually. Their approach involved a pre-processing pipeline, transfer learning, and fine-tuning the Pre-trained models. The methods encompassed pre-processing image pipelines for hair removal, dataset augmentation, resizing, transfer learning, and fine-tuning CNNs. EfficientNet B4 achieved an 87% F1 Score and 87.91% Top-1 Accuracy, with intermediate complexity models like B4 and B5 exhibiting superior performance. However, limitations included varying batch sizes due to computational constraints and the absence of universal standards for evaluating classification model performance.

Pallabi Sharma et al [4] the study highlighted centres on melanoma detection using advanced deep neural network models. It proposes adversarial training to enhance accuracy with limited annotated data, achieving notable success with ResNet101 at 84.77% accuracy. Methods employed included adversarial training, comparative analysis of pre-trained models like VGG16, VGG19, Densenet121, and Resnet101, and the utilization of GAN for automatic classification of melanoma skin cancer.

Results showcased ResNet101's achievement of 84.77% accuracy in melanoma classification, demonstrating the efficacy of adversarial training in enhancing accuracy with limited data. However, limitations persisted due to the restricted availability of annotated skin cancer images for training.

Aqsa Saeed Qureshi and Teemu Roos [5] focuses on skin cancer detection through an ensemble-based CNN architecture. Their proposed method enhances skin cancer detection accuracy, as evidenced by superior F1, AUC-PR, and AUC-ROC values compared to seven benchmark methods. The study's contributions include demonstrating these benefits using a dataset comprising dermoscopic images. However, limitations stem from the limited availability of qualified dermatologists and variations in dataset characteristics affecting method performance. The ensemble-based CNN architecture, leveraging pre-trained and trained models, showcases promise in advancing skin cancer detection methodologies.

Gourav Ganesh Ganesh and K Somasundaram Somasundaram [6] focuses on melanoma skin cancer detection via a deep-learning CNN model. They propose an enhanced CNN model with reduced computational cost, comparing it with pre-trained models like ResNet and DenseNet. Superior results were achieved by utilizing the HAM10000 dataset. Methods involve customizing the CNN model, adjusting activation functions and network architecture, and leveraging pre-trained models. Strategies for handling data uncertainties, including random sampling and noise reduction, were applied. Results show that the proposed CNN model outperforms pre-trained counterparts, highlighting its efficacy in melanoma detection using the HAM10000 dataset.

Hediye Orhan and Emrehan Yavsan introduces [7] an AI model for skin cancer diagnosis using deep learning techniques, contributing a model leveraging algorithms like AlexNet, MobileNet, ResNet, VGG16, and VGG19. Their proposed model developed on a dataset of 8,598 images, showcased superior performance. Notably, MobileNet achieved 84.94% accuracy. Which involve integrating the model with a desktop application and envisioning developing a mobile application for melanoma detection at home. The study's ethical standards were adhered to, with no human participants or animals involved, underscoring the model's potential for aiding melanoma diagnosis.

H.L. Gururaj et al [8] explores skin cancer detection using CNN models with the MNIST: HAM10000 dataset, contributing to early detection with models like DenseNet169. Data pre-processing techniques such as sampling, dull razor, and segmentation were employed for practical model training and transfer learning with ResNet50. Results showcased the utilization of DenseNet169 and ResNet50 for model training, alongside data pre-processing. This study underscores the significance of early skin cancer detection through the application of CNN models, supported by meticulous data pre-processing and transfer learning methodologies.

N. Priyadharshini et al [9] the study has introduced a new method for detecting skin cancer, specifically focusing on early detection of melanoma to improve treatment outcomes. The study proposes a hybrid algorithm, which combines the Extreme Learning Machine (ELM) and Teaching-Learning-Based Optimization (TLBO) algorithm, to accurately detect melanoma with higher precision and in less time than other methods. The study evaluated the accuracy of the proposed algorithm for disease prediction, as well as its computation time. The

results demonstrate that the ELM-TLBO algorithm has significantly improved melanoma detection accuracy. The study explains about the experimental findings and future directions for further research.

Krishna Mridha et al. [10] focus on applying deep learning techniques to improve the precision and efficacy of skin cancer diagnosis in order to reduce the workload for medical personnel. They have addressed concerns of class imbalance, developed deep learning models for the categorization of skin cancer, and proposed a novel end-to-end healthcare aid system using a mobile application. The work develops an explainable artificial intelligence (XAI)-based skin lesion categorization system for explanations and evaluates the efficacy of deep learning approaches with six classifiers. Techniques include experimenting with different activation functions and optimising CNNs (convolutional neural networks) utilising Adam and RMSprop functions. As evidenced by the results, which show a low loss accuracy of 0.47% and a high classification accuracy of 82%, the suggested approach is helpful in diagnosing skin cancer.

Maya V Karki and Santosh inamdar [11] explores skin cancer classification employing deep networks for precise diagnosis. The study employs transfer learning with pre-trained models like VGG16, VGG19, and ResNet50 on dermoscopy skin images. Their contributions encompass proposing an algorithm utilizing transfer learning with these models. The methods involve transfer learning, which utilizes knowledge from pre-trained models, and a technique called black top-hat filtering for hair removal to improve classification accuracy. The study emphasizes the significance of hair removal and deep networks in skin cancer classification.

Jinen Daghrir et al [12] study focused on melanoma detection through a fusion of deep learning and machine learning methodologies. Focusing on CNN utilization for effective melanoma lesion detection in skin images, the paper contributes by automating melanoma lesion detection and exploiting CNN efficiency. The research uses techniques like hair removal and lesion segmentation to leverage a dataset comprising 640 skin lesion images. Methodologies encompass deep learning with CNN, hairline detection via 2-D derivatives of Gaussian, and lesion segmentation through thresholding and morphological dilation. The study concludes that combining deep learning and classical ML techniques augments melanoma skin cancer detection.

Mahamudul Hasan et al [13] developed a skin cancer detection system utilizing Convolutional Neural Networks (CNN) for early diagnosis and treatment. The system attained an impressive 89.5% accuracy, evaluated via precision, recall, specificity, and F1 score metrics. Leveraging techniques such as image processing, feature extraction, and CNN classification, the study showcased promising results with 89.5% accuracy and 93.7% training accuracy. This research shows major contributions obtained in skin cancer detection, offering a valuable resource for early diagnosis and treatment strategies.

The study by Titus J. Brinker et al [14] compares the classification accuracy of artificial intelligence (AI) and dermatologists in identifying melanoma images. AI exhibited superior precision in this task compared to dermatologists, underscoring its potential in dermatology. Leveraging Convolutional Neural Networks (CNN), the research trained on 4204 biopsy-proven melanoma and nevi images, incorporating innovative deep learning techniques for refined classification. Results revealed that deep neural networks outperformed dermatologists in accurately classifying melanoma images. Furthermore, the study's integration of novel deep learning techniques further enhanced classification, signifying the promising role of artificial intelligence in advancing dermatological diagnostics.

Maria Colomba Comes et al [15] developed a deep learning algorithm aimed at forecasting disease-free survival among individuals diagnosed with cutaneous melanoma, aiming to optimize personalized treatments for aggressive forms of melanoma. The study addressed the critical need for reliable prognostic indicators in melanoma patient management by extracting prognostic biomarkers from whole-slide histological images. Leveraging deep learning on whole-slide images, the researchers extracted high-dimensional features for tasks, to get promising results. The model successfully predicted 1-year disease-free survival, achieving a median AUC of 69.5% and an accuracy of 72.7%. Validation cohorts confirmed these results, underscoring the potential of deep learning to enhance prognostic capabilities in melanoma patient care.

A study on melanoma detection was conducted by Sarah Ali Alshawi et al [16] the study aimed to enhance classification accuracy by using ensemble classifier models. They utilized ensemble transfer learning approaches to address identification errors, overfitting, and imbalanced classification. The researchers implemented classifiers like Adaboost, random forest, voted CNN and Boosted SVM, which resulted in high classification accuracy. The study demonstrated the effectiveness of ensemble classifier models and transfer learning methods in mitigating imbalanced classification issues. Evaluation of 19 classifiers through standard performance metrics emphasized the efficacy of these approaches in improving melanoma detection accuracy and handling classification challenges.

3. Proposed Methodology

3.1. Overall Workflow

Fig.1. depicts the overall workflow for the study's detection and classification of skin lesions using deep learning techniques. The HAM10000 dataset, which includes 10,015 different skin lesion photos, is initially prepared for analysis. The dataset is subsequently prepared for model training using image pre-processing techniques such as data augmentation and splitting. The pre-trained models, which include VGG-16, Inception V3, Inception ResNet V2, and DenseNet201, are then loaded for further processing. Model training begins with the creation of an initial CNN architecture, followed by fine-tuning the pre-trained models using techniques like transfer learning to improve performance. Evaluation metrics, including accuracy, are utilized to assess the effectiveness of the trained models.

Additionally, ensemble learning techniques combine fully fine-tuned Inception V3 and DenseNet 201 models to enhance classification accuracy further. Finally, a user-friendly web application is developed using HTML, CSS, JavaScript, and Python, providing an intuitive interface for skin lesion classification.

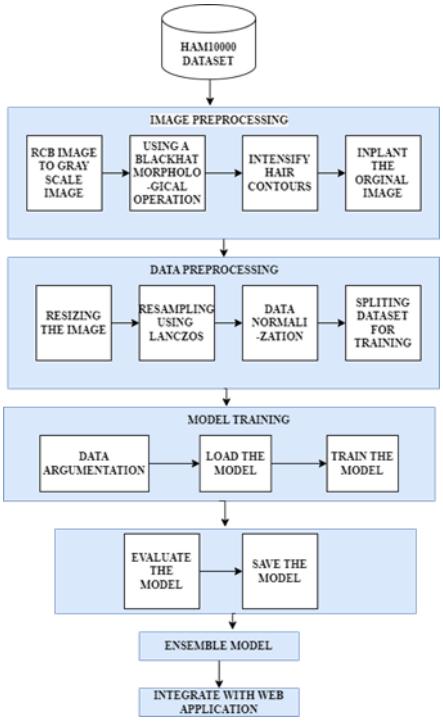

Fig. 1. Proposed Scheme workflow

3.2. Dataset Description

The HAM10000 training set comprises 10,015 dermatoscopic images. [17] These images represent pigmented lesions from diverse populations, with Austrian patients often referred to a specialized European center for early melanoma detection. In contrast, Australian patients typically exhibit severe chronic sun damage due to their residence in high skin cancer incidence areas.

Lesions in chronic sun-damaged skin often display multiple solar lentigines and ectatic vessels, with the occasional presence of small angiomas and seborrheic keratosis around the centrally located target lesion. Various imaging devices and dermatoscopic techniques, including polarized and non-polarized methods, were used to capture these images, ensuring a comprehensive representation of practical clinical scenarios.

The dataset covers 95% of lesions seen in clinical practice, aiding clinicians in making accurate diagnoses crucial for treatment. Lesions vary from benign to potentially cancerous, excluding non-pigmented variants. Of the 10,000 examples, 8,000 are benign, and 2,000 are malignant, with a focus on melanocytic nevi, comprising around 7,000 instances. Melanoma poses severe health risks, while basal cell carcinoma and actinic keratosis can progress to cancerous conditions.

Notably, the dataset is skewed towards melanocytic nevi, with nearly 7,000 examples can be seen in Fig. 2. The types of skin lesions present in dataset and their description can be seen in Table 1, below.

Table 1. Dataset Description

Class	Number of samples	% of class samples
Actinic Keratoses	327	3.27%
Basal Cell Carcinoma	514	5.13%
Begnign keratosis	1099	10.97%
Dermatofibroma lesions	115	1.15%
Melanocytic Nevi	6705	66.95%
Malignant Melanoma	1113	11.11%
vascular skin lesions	142	1.42%

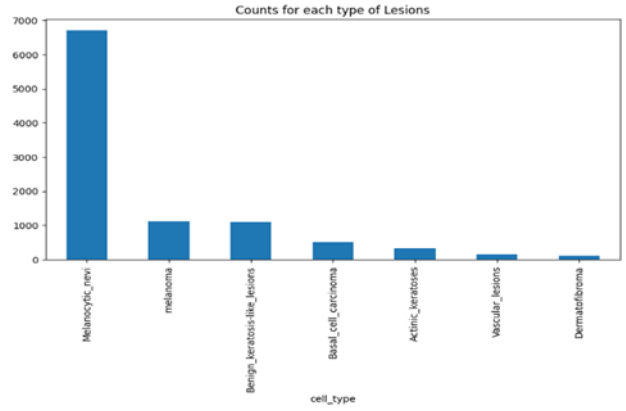

Fig. 2. Distribution of skin lesion categories.

3.3. Image Pre-Processing

The image pre-processing phase is essential in this study, primarily for removing unwanted elements and emphasizing the cancerous parts for accurate detection. Initially, RGB cancer images from the dataset are transformed into grayscale images to maintain crucial intensity information necessary for digital analysis [18]. Subsequently, a median filter is used to eliminate unwanted noise and enhance the overall image quality [19]. Detecting hair presence on the skin surface presents a notable challenge in cancer detection, necessitating additional processing steps. To address this challenge, hairy images are identified through a bottom hat filter, followed by the application of a specialized hair removal technique to effectively eliminate the detected hair portions from the images.

Black Hat-Top filtering, a technique widely used in morphology and digital image processing, plays a crucial role in extracting intricate components and details from images. It encompasses two fundamental transforms: the top hat and the black hat. The top-hat transform discerns the disparity between the input image and its opening, while the black-hat transform distinguishes the difference between the closing and the original image. These transformations are pivotal in various image-processing tasks, including feature extraction, background elimination, and image enhancement.

A novel technique has been introduced to tackle the issue of identifying and removing hair from dermoscopic images, enhancing the accuracy of classifying skin lesions. Unlike current methods, this innovative approach offers a simpler and more economical solution. By combining image filtering techniques with morphological operations applied to each input image, significant improvements can be made. The proposed method involves converting colour areas, performing basic kernel operations, applying morphological closing, and using image inpainting to reconstruct skin images that include hair.

The dataset utilized in this study comprises a diverse array of images, thereby complicating the detection process. However, by effectively addressing hair and various artifacts such as rulers and markers, optical diagnosis is streamlined. These procedures are meticulously executed through a series of steps.

3.3.1. Conversion of RGB Images into grayscale image

RGB images are converted to grayscale, emphasizing the red channel for its crucial information. This process condenses the image from three channels to one as shown in Fig. 3.

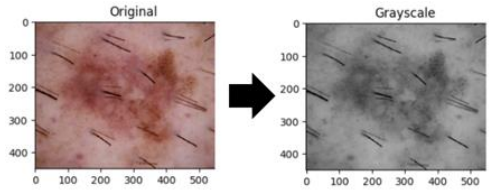

Fig. 3. Conversion of RGB image into grayscale image.

3.3.2. Identification of hairy regions

Hair regions are identified using a black-hat morphological operation, emphasizing larger and darker visual features compared to the surroundings. This technique employs morphological treatments, including gap, closing, erosion, and dilation, with a structuring size of 17 * 17 to detect hair contours as depicted in Fig. 4.

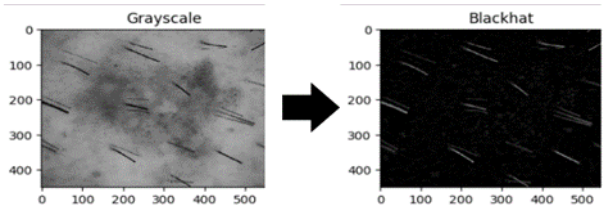

Fig. 4. Identification of hairy regions.

3.3.3. Enhancement of the contoured hairs

The intensity of hair contour images is enhanced by applying binary thresholding to increase the grayscale intensity of detected hair areas, as represented in Fig .5. By the following Equation (1).

$$I(x,y) = \begin{cases} 1 & if\ I(x,y) > threshold \\ 0 & otherwise \end{cases} \quad (1)$$

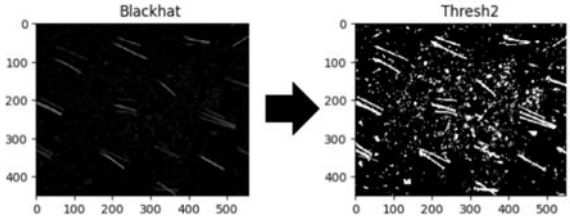

Fig. 5. Enhancement of the contoured hairs.

3.3.4. Detection of dark corners

Skin lesion images often exhibit dark, rounded edges, which may have similar pixel intensities to skin lesions themselves, potentially impacting categorization accuracy. These dark corners are identified and masked using Otsu thresholding. This technique optimizes threshold selection by iterating over various thresholds to minimize variance between foreground and background classes, as depicted in Equation (2).

$$\sigma_w{}^2(\mathrm{t}) = \sum_{i=1}^{t} \mathrm{p(i)(t)}\ \sigma_0{}^2\,(t) + \sum_{i=t+1}^{L} \mathrm{p(i)}\ \sigma_1{}^2\,(t) \quad (2)$$

3.3.5. Image inpainting

Total variation inpainting is employed to restore images obscured by hair. This technique is extended to 3D images, encompassing RGB data, to enhance skin lesion visibility [20]. By leveraging the original image's hair contours, the recovered image is obtained, as illustrated in Fig. 6. Using the equation (3).

$$V(I) = \sum_{i,j,k}^{\infty} [|I_{i+1,j,k} - I_{i,j,k}|^2 + |I_{i,j+1,k} - I_{i,j,k}|^2 + |I_{i,j,k+1} - I_{i,j,k}|^2]^{\frac{1}{2}} \quad (3)$$

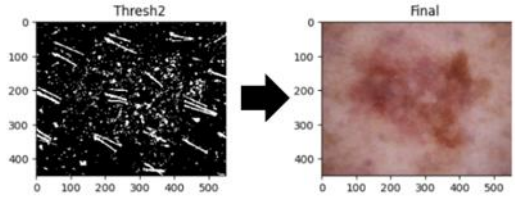

Fig. 6. Image inpainting.

3.4. Data Pre-processing

In the data pre-processing phase, the hair-removed images transform into a format suitable for training machine learning models. This involves resizing the images to a standard size and performing normalization to ensure consistent model input. The following steps outline the data pre-processing pipeline.

3.4.1. Image resizing

The hair images are resized to fit the input requirements of the models. For the baseline Convolutional Neural Network (CNN) model, images are resized to 64x64 pixels using Lanczos resampling. Similarly, images are resized to 256x192 pixels for pre-trained models using the same resampling technique.

3.4.1.1 Lanczos resampling

Lanczos resampling is a high-quality method for resizing images while preserving details and reducing artifacts. It is based on a convolution kernel defined by the Lanczos window function. The resampling function is represented by the following equation (4).

$$L(x) = \begin{cases} \sin(\pi x).\sin(\frac{\pi x}{a})/(\pi^2 x^2), & if\ -a < x < a \\ 0, & otherwise \end{cases} \quad (4)$$

Where a is the filter radius parameter, typically set to 3, Lanczos resampling applies this function to each pixel in the original image to compute its new value in the resized image.

3.4.2. Data normalization

After resizing, the pixel values of the images are normalized to enhance model convergence and stability during training. Two different normalization techniques are employed based on the type of model being used.

3.4.2.1 Data normalization for Baseline CNN model

The mean (μ) and standard deviation (σ) of the pixel values across all images are calculated. The pixel values are then normalized using the following equation (5)

$$X_{normalized} = \frac{X - \mu}{\sigma} \quad (5)$$

3.4.2.2 Data normalization for Pre-trained models

The pixel values are first converted to the float32 data type to ensure compatibility with the model architecture. Subsequently, the pixel values are scaled to the range [0, 1] using the following equation (6).

$$X_{scaled} = \frac{X}{255} \quad (6)$$

3.5. Data augmentation

In medical image analysis, where annotated datasets are often scarce and acquiring new information can be resource-intensive, data augmentation provides a powerful method to expand the available dataset and improve the performance of AI models. By generating diverse variations of existing medical images through transformations such as rotation, scaling, flipping, cropping, and more, data augmentation enables models to learn from a broader and more representative range of examples. Table 2 presents an overview of the data augmentation techniques utilized in this study, their descriptions and the corresponding parameters applied to train the model. This table facilitates an understanding of the specific image alterations applied and the impact of each technique on the augmentation process.

Table 2. Data augmentation applied on images

Technique	Description	Value	Action on the Image

Rotation	Rotates the image by a specified angle.	60	Counter clockwise rotation by the given angle.
Width Shift	Shifts the width of the image.	0.2	Horizontal shift by a fraction of the image width.
Height Shift	Shifts the height of the image.	0.2	Vertical shift by a fraction of the image height.
Shear	Applies shear transformation to the image.	0.2	Shearing along one axis while keeping the other axis fixed.
Zoom	Zooms the image in or out.	0.2	Zoom in or out of the image by the specified factor.
Fill Mode	Specifies how to fill the empty space created by the transformation.	Nearest	Filling empty space with the nearest pixel value.

3.6. Splitting of the data

Data splitting is a standard general method in most machine learning models; it helps to split a dataset into training and testing sets. Here, the images and labels are divided into training and testing sets, with test size=0.1, meaning 10% of the data is used for testing. Table 3 shows the shapes and descriptions of the arrays resulting from the splitting of the images as shown in Fig. 7.

Table 3. Data augmentation applied on images

Splitting of Data	Shape	Description
X_{train}	$(1-t)\times n$	Input features for training set from dataset
X_{test}	$t\times n$	Input features for testing set from dataset
y_{train}	$(1-t)\times 1$	Labels for training set from dataset
y_{test}	$t\times 1$	Labels for testing set from dataset

Where

$X_{train} \in R(1-t)\times n$ (Training set features)

$X_{test} \in Rt\times n$ (Testing set features)

$y_{train} \in R(1-t)\times 1$ (Training set labels)

$y_{test} \in Rt\times 1$ (Testing set labels)

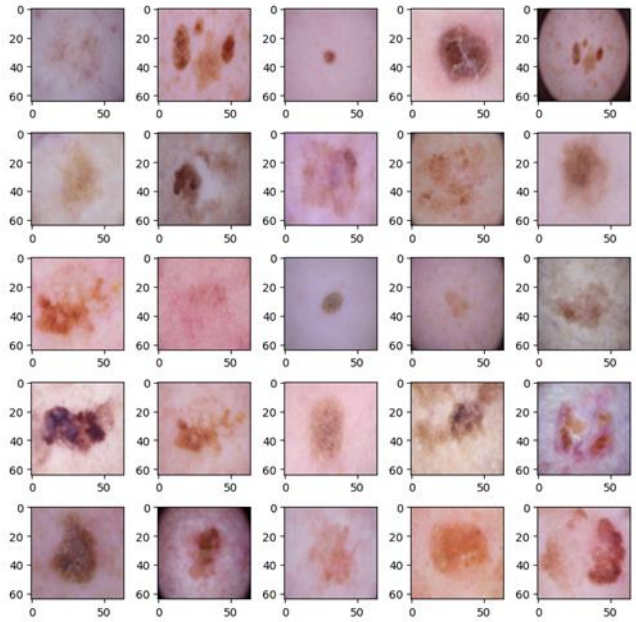

Fig. 7. Images given for training the models.

3.7. Model building

Our study uses the Keras library for model construction, leveraging its robust functionalities for deep learning tasks. From Keras, Proposed study import essential modules such as Dense, GlobalAveragePooling2D, and Dropout, complementing our model architecture. Additionally, we explore and assess the performance of several pre-trained models, including VGG16, DenseNet 201, Inception V3, and Inception ResNet V2. These pre-trained models have undergone extensive training on huge datasets like ImageNet, imparting them with valuable feature extraction capabilities. We fine-tune these models using our dataset for our skin lesion detection task, aiming to achieve optimal accuracy. The weights of these pre-trained models are imported from ImageNet, ensuring the transfer of learned representations for effective feature extraction.

Widely employed across various domains, including image classification, object detection, and segmentation, these pre-trained models offer versatility and robust performance, making them suitable candidates for our research.

3.7.1. Baseline Model

Before fine-tuning deep convolutional neural networks (DCNNs), a small CNN was constructed as shown in Fig. 8. The CNN architecture is based on heuristics, adhering to conventions observed in well-known DCNNs. The architecture incorporates three convolutional layers, with each followed by max-pooling layers aimed at decreasing the spatial activation size. The initial convolutional layer utilizes 16 kernels of size three, employing padding to preserve the image size. This is followed by a max-pooling layer with a 2x2 window, which shrinks the feature map by a factor of two. In the second convolutional layer, 32 kernels of size 3 are used, with padding to preserve spatial dimensions. The feature map size is then reduced further by applying another max-pooling layer. The third convolutional layer consists of 64 kernels of size 3, with padding to preserve spatial dimensions, followed by a final max-pooling layer. Data augmentation is employed during model training to introduce variation and ensure model generalization. The training begins with the Adam optimizer and a learning rate of 0.01. To optimise training, the learning rate lowers if the validation accuracy is constant for three epochs in a row. Each time this happens, the learning rate is cut in half. The training method takes 35 epochs in total.

The output Y (7) of a convolutional layer is calculated by convolving the input X with a set of filters W, followed by adding a bias term b, and applying an activation function f:

$$y = f((X * W) + b) \quad (7)$$

Max pooling down samples the input tensor by selecting the maximum value from each region of the input using equation (8).

$$Y_{i,j} = max_{m,n}(X_{i+m,j+n}) \quad (8)$$

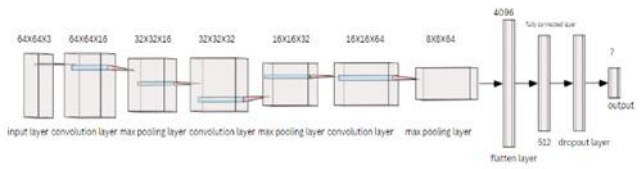

Fig. 8. Baseline CNN architecture.

3.7.2. VGG-16

Despite the availability of newer and more sophisticated DCNN models, VGG16 is chosen for fine-tuning due to its simplicity and ease of implementation. The VGG16 architecture shown Fig. 9. Featuring modified convolutional blocks with four convolutional layers in the third, fourth, and fifth blocks, achieves impressive accuracies on ImageNet: 90.1% for top-5 classifications and 71.3% for top-1 classifications.

For our skin lesion classification task, the top fully connected layers of VGG16 are replaced with custom layers tailored to the specific task. These include a global max-pooling layer, a fully connected layer with 512 units, a dropout layer with a rate of 0.5, and a softmax activation layer for classifying seven types of skin lesions.

To ensure effective fine-tuning, the initial training phase involves freezing all layers of VGG16 for feature extraction, followed by unfreezing the final convolutional block for fine-tuning. Training occurs over 20 epochs after three epochs of feature extraction, utilizing a learning rate of 0.001 and the Adam optimizer. During training, similar to the baseline model, strategies such as data augmentation and learning rate decay are employed. The VGG16 model is fine-tuned for a total of 30 epochs to better suit the skin lesion classification task.

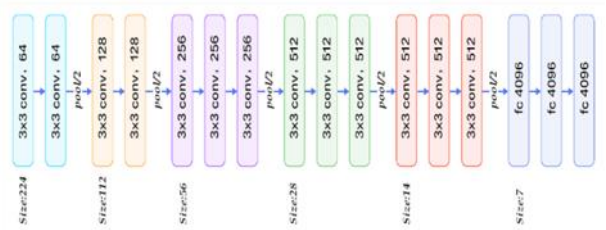

Fig. 9. VGG-16 architecture.

3.7.3. Inception V3

Inception V3, renowned for its outstanding performance on ImageNet with accuracies of 93.7% for top-5 classifications and 77.9% for top-1 classifications, features Inception modules that enable decision-making at each layer regarding the convolution type to apply. These modules encompass various convolution sizes, allowing the model to dynamically select the most suitable option. Additionally, Inception architecture efficiently facilitates the extraction of local and high-level features.

In the experimentation with Inception V3, two approaches were tested: fine-tuning from the last two inception blocks and fine-tuning the entire pre-trained model. This was motivated by considerations of Keras's Batch Normalization implementation. Keras's Batch Normalization implementation utilizes mini-batch statistics during training and previously learned statistics during inference, potentially leading to mismatched scaling between training and inference data if only the top layers are fine-tuned.

To tackle this issue, two fine-tuning strategies were investigated: fine-tuning all layers of Inception V3, and

fine-tuning the top two inception blocks while setting all Batch Normalization layers to trainable. Both approaches were executed for 35 and 20 epochs, respectively.

Inception-ResNet, another top performer on ImageNet, incorporates residual connections essential for training deep convolutional models. Fine-tuning strategies similar to those employed for Inception V3 were applied to Inception-ResNet V2, with the top layers fine-tuned for 30 epochs.

Batch normalization normalizes the activations of a layer by subtracting the batch mean μ and dividing by the batch standard deviation σ followed by scaling by learned parameters ϒ and β (9):

$$X' = \frac{X - \mu}{\sqrt{\sigma^2 + \varepsilon}}$$

$$Y = \Upsilon . X' + \beta \quad (9)$$

3.7.4. DenseNet

DenseNet, a novel DCNN architecture, has emerged as a top performer on ImageNet with impressive accuracies of 0.936 for top-1 and 0.773 for top-5 classifications. Compared to Inception V3, DenseNet achieves competitive performance with fewer parameters, approximately 20 million compared to Inception V3's 23 million. DenseNet 201, a variant of DenseNet, comprises four dense blocks, each characterized by a unique architecture.

Within a dense block depicted as Fig. 10, each layer receives inputs from all preceding layers and passes its feature maps to all subsequent layers, promoting feature reuse and information flow throughout the network. Unlike ResNet, which sums features, DenseNet concatenates features, explicitly distinguishing between added and preserved information. Each layer within a dense block carries out a composite function that includes three main steps: batch normalization, ReLU activation, and a 3x3 convolution operation.

Two experiments were conducted on DenseNet 201: fine-tuning the last dense block, which consists of 32 layers, and fine-tuning the entire network. The same training strategy used in previous sections was applied to both approaches. This involves fine-tuning the top layers of DenseNet 201 where it spans 27 epochs while fine-tuning the entire DenseNet 201 model lasts for 20 epochs. During fine-tuning, weights are initialized from the original pre-trained models on ImageNet for both DenseNet 201 and Inception V3.

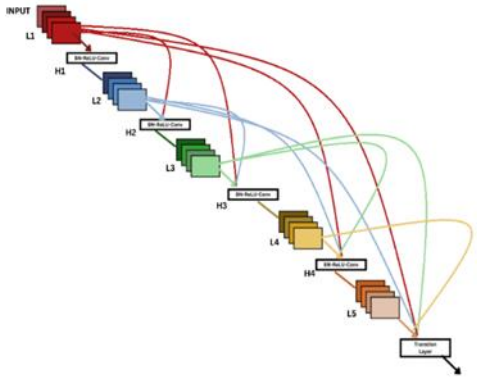

Fig. 10. Five Layer DenseNet Block.

3.7.5. Ensemble model

Constructed an ensemble comprising thoroughly fine-tuned Inception V3 and DenseNet 201 models using ensemble learning techniques. Leveraging the strengths of both models, the ensemble aimed to enhance classification performance.

3.7.6. Web Application

The major and final step left in the completion of this study is to utilize the Ensemble model and to integrate with an application which can be used by the society easily. An interactive, simple and user-friendly Web UI is created which can be seen as Fig.11 using HTML, CSS, JavaScript and Flask. The saved model is then integrated with the Flask framework. Each time the user uploads an image from the frontend UI, the CNN based classification model is called which extracts the lesion from the segmented image, parses through it and classifies it as Melanoma, Nevus, and Seborrheic Keratosis etc. The classification is then displayed along with the accuracy of prediction. Based on the convenience of the user, the app would also request for the user location and then suggest nearby hospitals and clinics which offer suitable treatment.

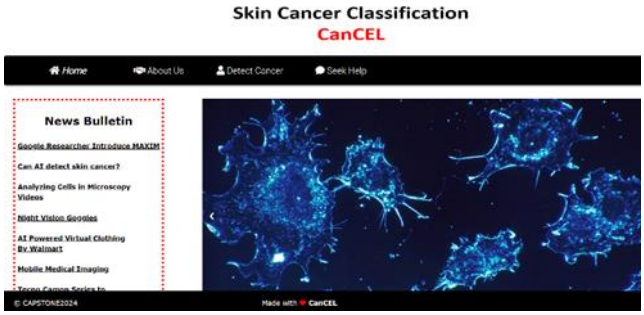

Fig. 11. Web application interface.

4. Experimental Results

In the experimental phase, VGG-16, Inception V3, Inception ResNet V2, and DenseNet 201 were subjected to fine-tuning. Initially, the top layers were frozen while the remaining layers were trained to adapt to the skin lesion classification task. By leveraging pre-trained

representations for lower-level features and enabling the adaptation of higher-level representations, the models underwent fine-tuning using a pre-processed dataset as depicted in Fig. 12.

4.1. Performance Metrics

Assessing the model is crucial to gauge its effectiveness and reliability in classifying skin cancer lesions. Accuracy serves as the primary metric to measure the model's performance.

Accuracy, as a metric, quantifies the overall correctness of the model by determining the ratio of correctly classified instances to the total number of instances.

$$Accuracy = \frac{Number\ of\ correctly\ classified\ instances}{Total\ Number\ of\ instances} \times 100$$

In Table 4, the results of the experimentation process are detailed, showcasing the validation accuracy, testing accuracy, and test loss for each model.

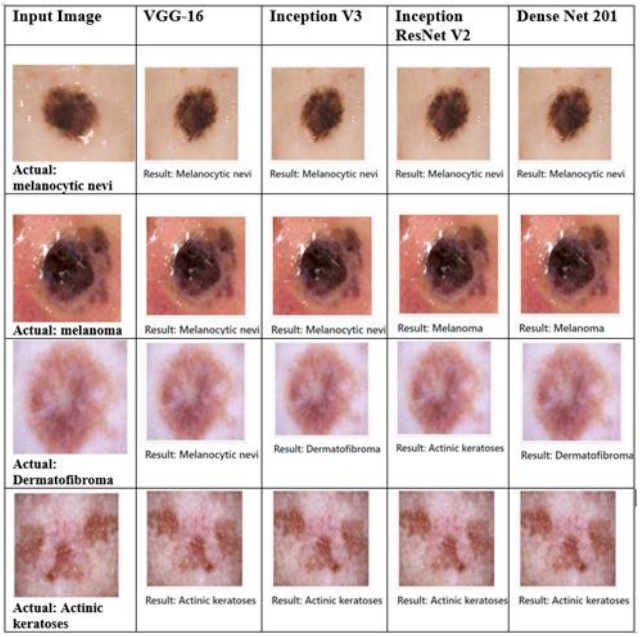

Fig. 12. Sample images for skin lesion classification.

Table 4. Fine tuning top layers of models

Model	Validation Accuracy	Testing Accuracy	Test Loss	Depth
Baseline Model	76.53%	75.64%	0.69	11 layers
VGG 16	79.83%	79.65%	0.71	23 layers
Inception V3	81.15%	82.33%	0.81	315 layers
Inception ResNet V2	83.03%	82.50%	0.82	784 layers
DenseNet 201	84.47%	85.02%	0.73	711 layers

5. Results

In the pursuit of accurate skin cancer detection, our research delved into addressing the challenging aspect of hair presence on the skin surface, which significantly impacts the reliability of detection algorithms. By implementing additional pre-processing steps, we aimed to mitigate this challenge and enhance the effectiveness of our detection model. Furthermore, leveraging the power of fine-tuning pre trained models, including VGG16, Inception V3, Inception ResNet V2, and DenseNet 201, enabled us to adapt these architectures to the intricacies of skin lesion classification. Through ensemble learning techniques, integrated the strengths of these models to achieve superior accuracies, as evidenced by the results showcased in Table 5. Our approach not only demonstrates the efficacy of our methodology but also underscores the importance of comprehensive pre-processing and model optimization strategies in advancing skin cancer detection technology.

Table 5. Fine tuning top all layers of models

Model	Validation Accuracy	Testing Accuracy	Test Loss
Inception V3	86.91%	86.83%	0.62
DenseNet 201	86.70%	87.72%	0.56
Ensemble Model	88.50%	88.52%	0.41

Table 6. Comparison between the proposed model and the states of arts

Study	Training algorithms	Dataset	Accuracy
A.A Nugroho, I. Slamet et al. [21]	CNN	HAM 10000	78%
Khan Adil & Iskandar et al. [19]	Alex Net	ISIC Dataset	84%
Hediya ORHAN, Emrehan YASAN. [7]	mobile Net	Kaggle Dataset (8598)	84.94%
K.Ali, Z.A shaikh et al. [3]	EfficentNet B4	HAM 10000	87.9%

Proposed model	InceptionV3 and DenseNet 201 (Ensemble Model)	HAM 10000	88.5%

By integrating meticulous image pre-processing, precise fine-tuning of pre-trained models, and strategic ensemble learning, our proposed methodology achieved an impressive accuracy of 88.5% than other other's work as shown in Table 6. This holistic approach optimized data quality, model architectures, and training strategies, resulting in superior performance in skin lesion classification.

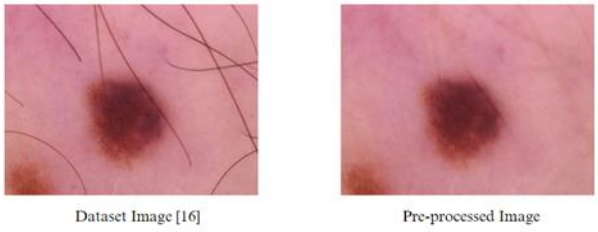

Fig. 13. Images used for proposed model training.

In our proposed study, pre-processed images were utilized for model training, contrasting with the approach of [18] and [7], who relied on the Ham10000 dataset for their research. This decision allowed us to optimize the quality and consistency of our training data, ensuring enhanced model performance and robustness as shown in Fig.14, Fig.15, Fig.16, and Fig17.

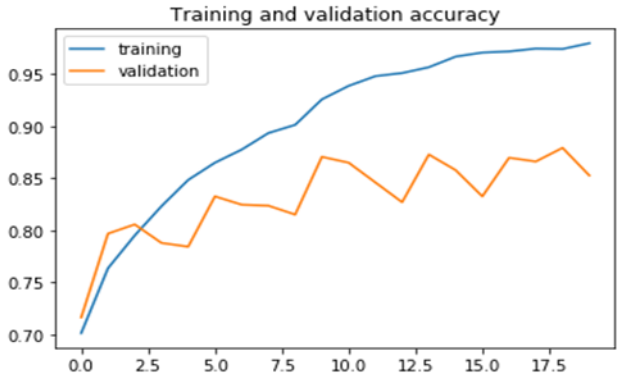

Fig. 14. Inception V3 Training and Validation accuracy graph.

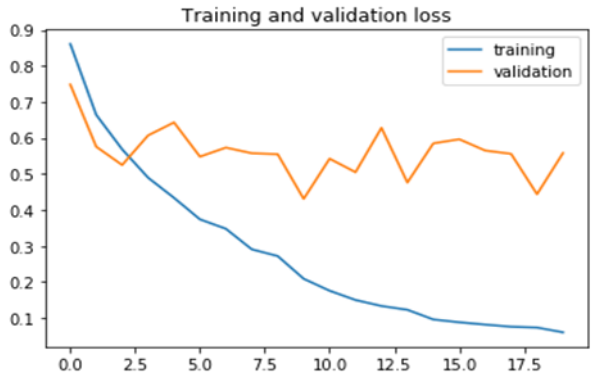

Fig. 15. Inception V3 Training and Validation loss graph.

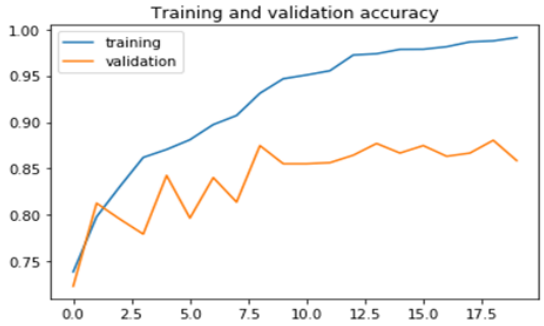

Fig. 16. DenseNet201 Training and Validation accuracy graph.

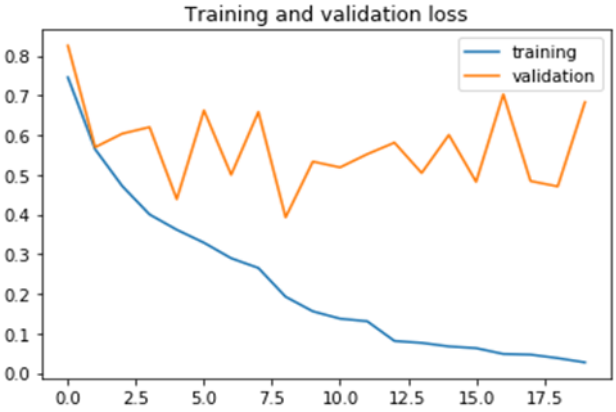

Fig. 17. DenseNet201 Training and Validation loss graph.

Fig. 18. Web application interface to detect skin lesion.

In the "Detect Cancer" section of our web application, shown in Fig. 18. Users can actively engage in the process of skin cancer detection by uploading photos of skin lesions. Leveraging the power of Flask, a Python micro-framework, we have seamlessly integrated our ensemble model into the backend, enabling real-time prediction of the lesion type based on the uploaded image. Through

sophisticated machine learning algorithms and ensemble learning techniques, our model accurately classifies skin lesions as depicted in Fig. 19, Fig. 20, Fig. 21, Fig. 22, Fig. 23, Fig. 24, and Fig. 25. Providing users with valuable insights into potential health risks. The user-friendly interface ensures a seamless experience, guiding users through the process step by step and presenting the prediction results in a clear and understandable format. By democratizing access to advanced diagnostic tools, our web application empowers individuals to take proactive steps towards early detection and treatment of skin cancer, ultimately contributing to improved health outcomes and well-being.

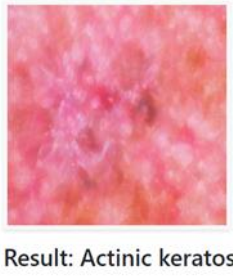

Fig. 19. Actinic keratoses detection.

Fig. 20. Basal cell carcinoma detection.

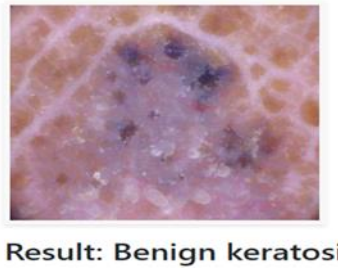

Fig. 21. Benign keratosis detection.

Fig. 22. Dermatofibroma detection.

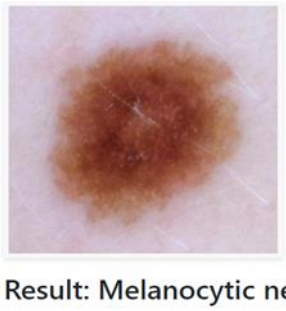

Fig. 23. Melanocytic nevi detection.

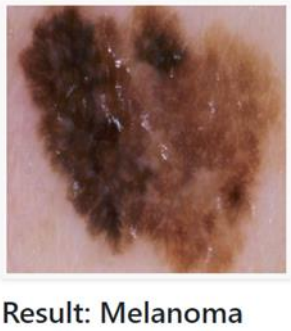

Fig. 24. Melanoma detection.

Fig. 25. Vascular lesion detection.

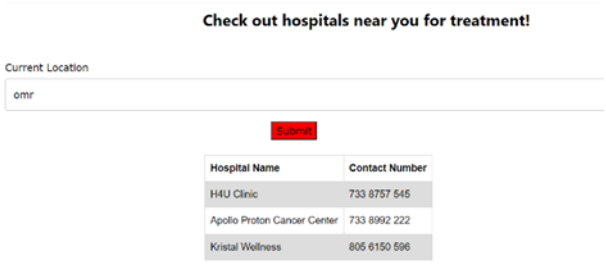

Fig. 26. Hospitals recommendation using user location.

In the "Seek Help" section of our web application, shown in Fig.26. Users can effortlessly locate nearby hospitals and medical facilities by simply inputting their location. Utilizing the capabilities of JavaScript, we've integrated dynamic functionality that retrieves and displays relevant hospital information based on the user's provided location. This feature enhances accessibility to essential healthcare services, facilitating prompt medical attention for individuals concerned about skin cancer or seeking

professional advice. By harnessing the power of technology to bridge the gap between users and healthcare resources, our web application empowers individuals to make informed decisions about their health and well-being.

6. Conclusion and Future Work

Fine-tuning all layers of the models proved to be a superior strategy compared to fine-tuning only the top layers, yielding better outcomes while requiring less time. Surprisingly, fine-tuning all layers was accomplished in a shorter duration than fine-tuning just the top layers, owing to the shorter epoch duration of 20 epochs compared to 30 epochs for the latter. This observation highlights the efficiency and effectiveness of fine-tuning the entire model, enhancing final performance and expediting the convergence process.

Future work addresses persistent overfitting issues, with current experiments showing 10-13% overfitting on training data despite attempted mitigation strategies. Further research into advanced methods for overfitting reduction is imperative to enhance model generalization and overall performance. Additionally, expanding the project to develop an end-to-end application for skin lesion detection and management is recommended. This application could include features for storing patient details, enabling comprehensive patient management and longitudinal monitoring of skin conditions. Such efforts aim to improve healthcare outcomes by facilitating early detection, accurate diagnosis, and personalized treatment planning for skin lesions.

Author contributions

Janardhan Kantubhukta: Investigation, Conceptualization, Writing-original draft. **Dr.Madheswari K:** Review and editing **Dr.Karthikeyan N:** Review and editing.

Conflicts of 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.

References

[1] Alavikunhu Panthakkan, S. M. Anzar, S. Jamal, and W. Mansoor, “Concatenated Xception-ResNet50 — A novel hybrid approach for accurate skin cancer prediction,” Computers in Biology and Medicine, vol. 150, pp. 106170–106170, Oct. 2022.

[2] D. K. A. AL-SAEDİ and S. SAVAŞ, “Classification of Skin Cancer with Deep Transfer Learning Method,” Computer Science, Sep. 2022.

[3] K. Ali, Z. A. Shaikh, A. A. Khan, and A. A. Laghari, “Multiclass skin cancer classification using EfficientNets – a first step towards preventing skin cancer,” Neuroscience Informatics, vol. 2, no. 4, p. 100034, Dec. 2022.

[4] P. Sharma, A. Gautam, R. Nayak and B. K. Balabantaray, "Melanoma Detection using Advanced Deep Neural Network," 2022 4th International Conference on Energy, Power and Environment (ICEPE), Shillong, India, 2022, pp. 1-5.

[5] A. S. Qureshi and T. Roos, “Transfer Learning with Ensembles of Deep Neural Networks for Skin Cancer Detection in Imbalanced Data Sets,” Neural Processing Letters, Oct. 2022.

[6] Ganesh, Gourav Ganesh, and K. Somasundaram Somasundaram. "Detect Melanoma Skin Cancer Using An Improved Deep Learning CNN Model With Improved Computational Costs." (2023).

[7] Hediye ORHAN and Emrehan YAVŞAN, “Artificial intelligence-assisted detection model for melanoma diagnosis using deep learning techniques,” Mathematical Modelling and Numerical Simulation with Applications, vol. 3, no. 2, pp. 159–169, Jun. 2023.

[8] H. L. Gururaj, N. Manju, A. Nagarjun, V. N. M. Aradhya and F. Flammini, "DeepSkin: A Deep Learning Approach for Skin Cancer Classification," in IEEE Access, vol. 11, pp. 50205-50214, 2023.

[9] N. Priyadharshini, Selvanathan N., B. Hemalatha, C. Sureshkumar, A novel hybrid Extreme Learning Machine and Teaching–Learning-Based Optimization algorithm for skin cancer detection, Healthcare Analytics, Volume 3,2023, 100161,ISSN 2772-4425.

[10] K. Mridha, M. M. Uddin, J. Shin, S. Khadka and M. F. Mridha, "An Interpretable Skin Cancer Classification Using Optimized Convolutional Neural Network for a Smart Healthcare System," in IEEE Access, vol. 11, pp. 41003-41018, 2023.

[11] M. V. Karki and S. Inamdar, "Skin Cancer Classification Using Deep Networks," 2022 4th International Conference on Circuits, Control, Communication and Computing (I4C), Bangalore, India, 2022.

[12] J. Daghrir, L. Tlig, M. Bouchouicha and M. Sayadi, "Melanoma skin cancer detection using deep learning and classical machine learning techniques: A hybrid approach," 2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), Sousse, Tunisia, 2020, pp. 1-5.

[13] Surajit Das Barman, M. Hasan, and F. Roy, “A Genre-Based Item-Item Collaborative Filtering,” Feb. 2019.

[14] T. J. Brinker et al., “Deep neural networks are superior to dermatologists in melanoma image classification,” European Journal of Cancer, vol. 119, pp. 11–17, Sep. 2019.

[15] M. C. Comes et al., “A deep learning model based on whole slide images to predict disease-free survival in cutaneous melanoma patients,” Scientific Reports, vol. 12, no. 1, Nov. 2022.

[16] Alshawi, Sarah Ali, and Ghazwan Fouad Kadhim AI Musawi. "Skin cancer image detection and classification by cnn based ensemble learning." International Journal of Advanced Computer Science and Applications 14, no. 5 (2023).

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

[18] Han, Seung Seog, et al. "Keratinocytic skin cancer detection on the face using region-based convolutional neural network." JAMA dermatology 156.1 (2020): 29- 37.

[19] H Khan, Adil, et al. "Classification of skin lesion with hair and artifacts removal using black-hat morphology and total variation." International Journal of Computing and Digital Systems 10 (2021): 597-604.

[20] T. Y. Satheesha, D. Satyanarayana, and M. N. Giriprasad, “A Pixel Interpolation Technique For Curved Hair Removal In Skin Images To Support Melanoma Detection,” J. Theor. Appl. Inf. Technol., vol. 31, no. 3, 2014.

[21] A. A. Nugroho, I. Slamet, and Sugiyanto, “Skins cancer identification system of HAMl0000 skin cancer dataset using convolutional neural network,” INTERNATIONAL CONFERENCE ON SCIENCE AND APPLIED SCIENCE (ICSAS) 2019.

[22] A. Khan, D. Iskandar, J. F. Al-Asad, and Samir El-Nakla, “Classification of Skin Lesion with Hair and Artifacts Removal using Black-hat Morphology and Total Variation,” International Journal of Computing and Digital Systems, vol. 10, no. 1, pp. 597–604, May 2021.