Karbala International Journal of Modern Science

Manuscript 3435

Skin Cancer Image Classification Using Deep Learning with Data Segmentation Technique

Akhilesh Kumar Shrivas

Hema Vastrakar

Skin Cancer Image Classification Using Deep Learning with Data Segmentation Technique

Abstract

The human skin is an impressive organ and structural element often impacted by a diverse range of recognized and unknown diseases. Diagnosing disorders that affect the outermost layer of the body is the most uncertain and difficult component in the scientific field. Dermatological diseases are one of the most significant health concerns in the 21st century since their identification is challenging and costly, plagued with challenges and the subjectivity that comes with human interpretation. The main objective of this piece of research work is to develop a robust model for the classification of skin cancer diseases using deep convolution neural networks (CNN) and transfer learning models. This research work uses convolutional neural networks(CNN), ResNet50, and VGG-16 techniques for the classifications of skin lesions. Data augmentation and UNet++ techniques are used to improve the quality of images and the performance of CNN, VGG16, and RestNet50 models. This research work is divided into three cases to analyze and evaluate the performance of models. In the first case, the deep learning and transfer learning models are used for the classification of skin cancer. In the second case, the data augmentation technique is applied to the skin cancer dataset and improves the model performance in terms of accuracy. In the third case, the deep learning and transfer learning models with data augmentation and UNet++ segmentation techniques are applied to develop a proposed model with high accuracy for the classification of skin cancer. The proposed VGG16-DAUNPP model achieved better 98.83% and 98.63% accuracy in the case of the Ham10000 dataset and Malignant vs. Benign ISIC dataset respectively. Similarly, the ResNet50-DAUNPP model achieved the highest 98.87% accuracy as compared to others in the case of Melanoma skin cancer dataset. Finally, our proposed transfer learning models VGG16-DAUNPP and ResNet50-DAUNPP are robust and recommended for the classification of skin cancer diseases.

Keywords

Skin Disease, Classification, Deep Learning, Transfer learning, Data Augmentation, Segmentation Technique

RESEARCH PAPER

Skin Cancer Image Classification Using Deep Learning with Data Segmentation Technique

Akhilesh K. Shrivas, Hema Vastrakar*

Department of CSIT, Guru Ghasidas Vishwavidyalaya, 495009, Bilaspur, Chhattisgarh, India

Abstract

The human skin is an impressive organ and structural element often impacted by a diverse range of recognized and unknown diseases. Diagnosing disorders that affect the outermost layer of the body is the most uncertain and difficult component in the scientific field. Dermatological diseases are one of the most significant health concerns in the 21st century since their identification is challenging and costly, plagued with challenges and the subjectivity that comes with human interpretation. The main objective of this piece of research work is to develop a robust model for the classification of skin cancer diseases using deep convolution neural networks (CNN) and transfer learning models. This research work uses convolutional neural networks (CNN), ResNet50, and VGG-16 techniques for the classifications of skin lesions. Data augmentation and UNet++ techniques are used to improve the quality of images and the performance of CNN, VGG16, and RestNet50 models. This research work is divided into three cases to analyze and evaluate the performance of models. In the first case, the deep learning and transfer learning models are used for the classification of skin cancer. In the second case, the data augmentation technique is applied to the skin cancer dataset and improves the model performance in terms of accuracy. In the third case, the deep learning and transfer learning models with data augmentation and UNet++ segmentation techniques are applied to develop a proposed model with high accuracy for the classification of skin cancer. The proposed VGG16-DAUNPP model achieved better 98.83 % and 98.63 % accuracy in the case of the Ham10000 dataset and Malignant vs. Benign ISIC dataset respectively. Similarly, the ResNet50-DAUNPP model achieved the highest 98.87 % accuracy as compared to others in the case of Melanoma skin cancer dataset. Finally, our proposed transfer learning models VGG16-DAUNPP and ResNet50-DAUNPP are robust and recommended for the classification of skin cancer diseases.

Keywords: Skin disease, Classification, Deep learning, Transfer learning, Data augmentation, Segmentation technique

1. Introduction

Cancer is a deadly disease that affects the lives of all individuals. In some situations, it could be harmful to humans. Although many different cancers can develop in the human body, skin cancer is infamous for its rapid growth and high mortality rate. Several things might trigger it, including smoking, alcohol use, allergies, infections, viruses, UV radiation exposure, physical exercise, environmental changes, and so on. Sunlight's ultraviolet (UV) rays can damage or even destroy the DNA within skin cells. One other thing that can cause skin cancer to develop is an abnormal enlargement of the human body. Skin cancer is one of the most common types of cancer in the world and is growing quickly. Early detection is essential for successful treatment and enhanced patient outcomes. Recent progress in artificial intelligence (AI) and deep learning (DL) has shown a lot of promise for automating the classification of skin lesions. This could help dermatologists make clinical diagnoses. But there are still problems like class imbalance, differences in how lesions look, and not enough expert-annotated datasets. Researchers have looked into methods including data augmentation, image segmentation, and model ensembles

Received 27 May 2025; revised 15 September 2025; accepted 18 September 2025.
Available online 15 October 2025

* Corresponding author.
E-mail addresses: proffshrivas@gmail.com (A.K. Shrivas), hema75hr@gmail.com (H. Vastrakar).

https://doi.org/10.33640/2405-609X.3435

to deal with these issues. Our research expands upon these advancements by examining an innovative integration of VGG16 with UNet++-based segmentation and data augmentation (DA) methodologies. We have tested our models on three publicly available datasets: HAM10000, ISIC2019, and a collection of melanoma images. This will help us see whether it works well on different types of images as well as if it is accurate. This work contributes to the ongoing efforts in AI-assisted dermatology by proposing a hybrid deep learning pipeline optimized for skin cancer classification. The most common types of skin cancer are melanoma, basal cell carcinoma, actinic keratoses, and squamous cell carcinoma [1]. Melanoma is a common type of cancer that happens when pigmented cells on the skin's surface grow out of control. This disease is quite bad and can spread swiftly to other parts of the body. It is most common in those with lighter skin. It consistently destroys its victims [2]. This illness can cause black lesions with an unusual border to appear naturally on otherwise healthy skin. Melanoma can also develop from a mole that has multiple colors, an unstructured development of the lesion, and an enlargement of the lesion location [3]. On the other hand, the diagnosis may be delayed at times, and the early discovery of a problem may not always cause the patient to be concerned. There is a possibility that the manual diagnostic will take an excessive amount of time, and the majority of patients are unable to pay for it. The incidence of skin cancer has increased, a matter of concern for many. Research conducted by experts indicates that there were 1.4 million new cancer cases in India in 2020, with projections suggesting an increase to 1.57 million cases by the end of 2025 [37]. This pace of increase is concerning to individuals, as it appears that the various groups within the nation are increasingly impacted. This trend indicates a higher prevalence of cancer among the population in the country. In 2025, the projected global incidence of new skin cancer cases is 107,240, including 59,080 instances of ductal carcinoma in situ among women, as per a report. The survey also forecasts around 2,041,910 new cancer cases globally in 2025. The precise number of skin cancer cases in India in 2025 is not detailed in the available search results; nonetheless, forecasts suggest a substantial rise in total cancer cases in India by 2025, with an anticipated increase of 12.8 % compared to 2020 [38].

The standard medical approach to skin disease diagnosis relies on the doctor's expertise, clinical experience, and dermatoscopic pictures to determine the severity of the patient's skin disease. One approach to think about the diagnostic process is that the doctor will first apply visual diagnosis to get information about the patient. After that, histological investigation and dermoscopy will be used. Dermoscopy is a high-definition imaging technology that doesn't irritate the skin. It may look into the skin's architecture where the surface dermis and inner dermis meet [25]. Unfortunately, only doctors with a lot of experience can tell the clinical signs of pigmented skin illnesses. This is because skin patches have similar colors, textures, edges, and other features, yet aberrant tissues from various patients are not the same. On the other hand, the patient doesn't need medical tools that are based on experience. In usual, it takes 4–5 days from the sample test to the doctor's evaluation, the histopathological research, and then the patient's report. This process takes a long time, which slows down the patient's rehabilitation. The following are some of the main problems with conventional dermatological diagnosis: Firstly, there is an insufficient availability of medical services. Professional dermatologists who have experience in skin care are in limited numbers. Many people are without a connection with dermatologists because the number of dermatologists has been unable to keep ahead with the rise in skin diseases. And secondly, the accuracy of diagnostics is low.

Although several studies have explored CNNs and transfer learning for skin cancer classification, most are limited to binary classification and do not incorporate ensemble architectures or segmentation techniques such as UNet++. Moreover, they rarely address the real-world clinical challenge of multi-class classification involving various lesion types, which is essential for practical deployment. Recent work such as the study in Ref. [41] highlights ongoing limitations in generalizability, data diversity, and high misclassification risks in underrepresented lesion types. Consequently, there is a requirement for the development of a deep learning model that is superior to human diagnosis, reduces the amount of time needed to generate results, and produces more accurate results [4]. Deep learning techniques have addressed problems in traditional skin cancer disease diagnostics and image recognition technology. Image recognition using deep learning is a multifaceted field. Using feature engineering and Deep learning-based classification algorithms, we integrate medical skin cancer disease imaging, mathematical methodologies, and computer technology to accurately identify and diagnose skin diseases. The picture recognition method for skin disease is founded on deep learning techniques used to

extract the characteristics of skin diseases. Conducting deep exploratory data analysis and reducing its dimensions demands extensive professional medical understanding [26]. Resulting in complex parameter adjustments might be time-consuming and energy-intensive. This approach is limited to a certain field due to its low mobility. Advancements in deep learning have led to successful picture recognition outcomes. Deep learning can automatically identify nonlinear relationships in medical images without the need for feature engineering, and extractive efficiency is also high. Deep learning technology is extremely customizable and may be applied to various fields and applications [27]. As shown in Fig. 1, the malignant form of skin cancer, such as melanoma, can be distinguished from benign skin lesions, such as melanocytic nevi, benign keratosis, dermatofibroma, actinic keratosis, and vascular lesions. Melanoma is an instance of a malignant form of skin cancer disease. Several distinct kinds of research projects as well as agendas [5,6] studies used convolutional neural networks, or CNNs, to identify skin cancer and classify images into melanoma and non-melanoma categories.

Therefore, the current research aims to answer the following questions like:

- Can a deep learning approach incorporating data augmentation, UNet++, and transfer learning significantly improve the accuracy of skin cancer classification?
- How effective are CNN, ResNet, and VGG-16 models in distinguishing between melanoma and other benign or malignant skin conditions in a multi-class setting?
- Can automated image-based diagnosis outperform manual methods in terms of speed and accuracy, particularly for early-stage detection?
- Is the use of UNet++ beneficial in segmenting lesion boundaries to enhance feature extraction for classification models?

This work mainly focused on improving the functionality of the present model by including data augmentation, UNet++, and transfer learning models into the observed dataset for multi-class classification, even though these approaches provide good results. This study proposes a hybrid diagnostic pipeline that combines segmentation (UNet++) and classification (CNN, VGG-16, ResNet) techniques to enhance detection of skin lesion types. Unlike many existing works, it emphasizes multi-class classification of skin lesions, better representing clinical diagnosis scenarios. It integrates data augmentation to counter data imbalance issues, and boosting model generalization. The workflow minimizes diagnostic delay while improving automated accuracy, aiming to assist in real-time clinical use. The ultimate objective of our proposed approach is to identify melanoma and other skin lesions through multi-class envision classification. Through the examination of pictures of skin lesions, our primary objective is to develop a robust model for identifying cases of skin cancer as well as other forms of skin diseases by employing the CNN algorithm, and ResNet, and VGG-16 transfer learning models with data augmentation, and UNet++ techniques.

The application of deep learning and transfer learning models such as VGG16 and ResNet50, alongside data augmentation and segmentation techniques like UNet++, has been widely used in the domain of skin cancer classification. However, most prior studies have explored these techniques in isolation or without systematically analyzing their integrated impact on performance across multiple real-world datasets. In this research, we introduce a comprehensive, empirically validated pipeline that combines transfer learning (VGG16 and ResNet50) with dual preprocessing mechanisms, advanced data augmentation and UNet++ segmentation technique to evaluate performance variations across three distinct public datasets: HAM10000, ISIC2019, and the Melanoma Skin

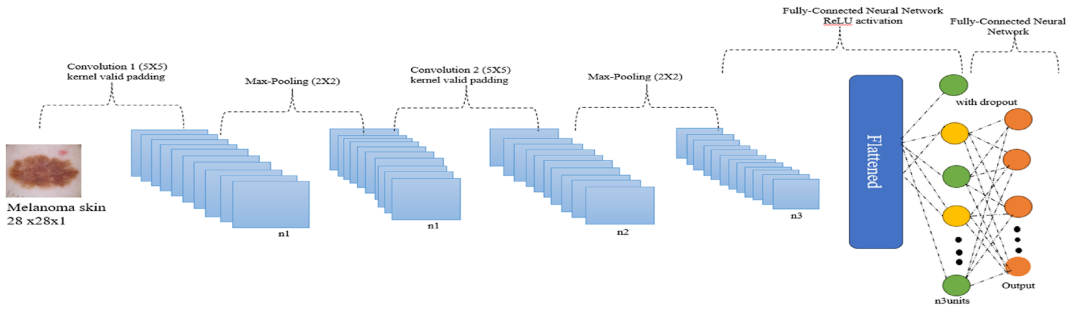

Fig. 1. Classification of skin lesions using CNN.

Cancer dataset. The study's novelty comes from its structured comparison framework across three experimental cases (Case 1: without DA and UNet++, Case 2: with DA only, and Case 3: with DA and UNet++), that helps in isolating the individual and combined effects of these preprocessing techniques. Furthermore, the study creates a uniform benchmarking platform, reports accuracy changes with and without augmentation/segmentation, and compares the final results to the most recent cutting-edge models. Unlike previous works, which frequently report single-model performance on isolated datasets, our approach focusses on the proposed hybrid pipeline's generalisability and scalability, proving superiority in both binary (malignant vs benign) and multi-class skin lesion classification tasks. This structured analysis methodology, together with the incorporation of UNet++ and DA into transfer learning-based pipelines, makes an important contribution to optimizing real-world skin cancer detection procedures.

2. Related work

Dermatologists spend an enormous amount of time on patient visits and conversations to manually recognize skin problems. The majority of the population lacked this option. A nearby metropolis is the best place for residents of these smaller villages to go for assistance and to file a report. So much effort is required of individuals for this. A visit to the doctor already adds up to a huge expense. Additionally, this involves interacting with other people, which is problematic given the current pandemic crisis. The transmission of diseases is rare. Presently, there is no way of avoiding physical touch. Most modern forms of computer-assisted diagnosis classify skin injuries and burns as diseases. We can do a better job using these methods. A computer-aided method that can detect skin-related conditions and discriminate them from other skin problems is consequently urgently needed. Q. Gan et al. [7] explored the image's color and structure to find diseases on the skin. Median filtering was used to get the pictures ready for processing. Images that have been cleaned up are turned to get their parts. We used the GLCM tool to pull out text features and the SVM tool to put skin diseases like herpes, eczema, and psoriasis into groups. P. N. Srinivasu et al. [8] explored MobileNet V2 and long short-term memory, which are built on deep learning, to classify skin diseases. To Fig. out how fast a sickness is spreading, a grey-level co-occurrence matrix was used. S. Malliga et al. [9] investigated CNN as a potential tool for learning and classifying a wide range of clinical images. Three new types of skin disorders have been identified. They were more complicated and go by names like melanoma, nevus, and seborrheic keratosis. N. Hameed et al. [10] efficiently implemented a method to categorize images of skin cancer dieas either benign, acne, eczema, healthy, or malignant melanoma. Utilizing Alex NET, a pre-trained CNN model, they were able to acquire the characteristics. The sorting was done using a support vector machine technique, which yielded somewhat more accurate outcomes in the end. M. A. Milton et al. [6] facilitated the detection of skin malignancies including melanoma. Using the ISIC 2018 dataset, the author tests several deep learning models, including PNASNet-5-large, Inception ResNetV2, SENet154, and InceptionV4. As an optimizer, adaptive momentum (Adam) was employed, and the models were trained using various parameters and learning rates. In the results, they assigned a class number to each skin cancer disease kind. After reviewing all of the models, the author gave the highest confirmation score to PNASNet-5-large. Even though it did a decent job of classifying different skin lesions, as this study demonstrated. T. Sreelatha et al. [12] implied pattern recognition techniques in image processing by numerous healthcare providers that can be employed for the early detection of skin cancer disease. The author developed a GFAC-based melanoma segmentation approach to detect early-stage melanoma skin cancer. The suggested method was used on the Ph2 dataset. This helped get rid of noise and made execution faster. Image segmentation, on the other hand, is more accurate than other classification methods. K. Matsunaga et al. [5]implemented the model with ISIC dataset to categorize seborrheic keratosis, nevi, and melanoma, among other skin cancer disease. Experiments made use of deep learning models that had previously been trained on ImageNet, ResNet-101, and InceptionV4. Using data augmentation, the images were also cropped, inverted, and expanded. M. Q. Khan et al. [13]outlined a comprehensive methodology for identifying and classifying melanoma and nevi in their article. The authors initiated the process by employing a Gaussian filter to eliminate the noise. The process of identifying and separating lesions involved using the technique of K-mean clustering. The color and texture features were subsequently extracted using a hybrid super-feature vector. Subsequently, support vector machines (SVMs) were utilized in the classification process. Implementing the suggested approach on the ERMIS dataset produced

enhanced accuracy. Recent studies have made significant progress in applying deep learning for skin cancer detection. For instance, D. Meedeniya et al. [41] proposed an effective deep learning system that integrates MobileNetV2 and Vision Transformers (ViT) to improve the categorisation of skin lesions, specifically melanoma, basal cell carcinoma, and benign keratosis, with exceptional accuracy with the HAM10000 dataset. They developed an attention-based technique to concentrate on features pertinent to lesions, enhancing classification performance and diminishing noise. There are certain problems with their study that our current research is trying to fix: Their classification is less clinically useful in real-world situations because it only covers three main types of lesions. Our work expands the classification to include more types of lesions, such as dermatofibroma, melanocytic nevi, actinic keratosis, and vascular lesions. Absence of lesion segmentation: Their model fails to employ explicit lesion border segmentation, which is essential for feature localisation. In contrast, our method uses UNet++, a sophisticated segmentation model, to accurately find lesion areas before classifying them. Model architecture: They employed MobileNetV2-ViT fusion, while we use a hybrid architecture that combines CNN, ResNet, VGG-16, and UNet++, coupled with data augmentation, to make the model work better on a wider range of samples. P.A. Diallo et al. [20] used several methods of data enhancement, including color manipulation, to bring forth greater detail in the photos. Resolution can also be increased with the help of Generative Adversarial Networks (GANs). Many methods of picture production employ GANs to improve image quality through the fusion of images. I. Pacal et al. [42] proposed a Convolutional Swin Transformer (CSwinformer) architecture specifically designed for the simultaneous segmentation and classification of skin lesions. The suggested model exploited hierarchical self-attention in shifted windows to improve both local lesion localisation and global context awareness. It did better than traditional CNNs in both accuracy and border detection. This was especially important for dermoscopic images, since the edges of lesions are often unclear or uneven. Their algorithm was tested on public datasets including HAM10000 and ISIC-2019, and it was better at telling the difference between melanoma and other types of skin cancer. G. H. Dagnaw et al. [43] implemented a comprehensive analysis of Swin Transformer (tiny/base versions), Vision Transformers (ViT), and CNN models including ResNet and MobileNet, incorporating explainable AI (XAI) methodologies such as CAM and Grad-CAM++. Their research focused on model interpretability, enhancing its applicability in clinical settings. XAI showed the parts of the lesion that helped with the classification decision, which is an important part of medical AI. A. Mahbod et al. [44] proposed a new hybrid architecture by combining PanDerm with Swin Transformer V2-base. This combination greatly improved performance on tough datasets like HAM10000 and MSKCC, especially when it came to telling apart classes that are hard to tell apart, such actinic keratosis and squamous cell carcinoma. Their findings validated that transformer-based vision backbones can surpass CNNs when combined with domain-specific pretrained models. J. Aftab et al. [45] explored the application of EfficientNet-B4 for multi-class skin lesion classification, attaining an F1-score of 87.3 % and an accuracy of 88.1 % on the ISIC-2020 dataset. The EfficientNet family, known for its compound scaling (depth, width, resolution), achieved an optimal balance between performance and computational cost. The authors demonstrated enhanced generalisation in under-represented lesion categories when compared to ResNet and Inception-v3. V. Venugopal et al. [46] implemented a customised EfficientNet variant to classify seven types of skin lesions, leveraging dermoscopic features including asymmetry, border irregularity, and colour variation. The model exhibited enhanced specificity and sensitivity, highlighting its effectiveness in both clinical-grade binary and multi-class scenarios. However, it lacked adequate lesion segmentation, which could undermine model transparency in real-world diagnostic applications. M.U. Ali et al. [47] developed a fusion ensemble model that combines EfficientNet, DenseNet, and Inception architectures to tackle dermatological conditions such as eczema, melanoma, basal cell carcinoma (BCC), and psoriasis. The ensemble model exhibited notable robustness and generalisation, particularly when evaluated on imbalanced datasets, highlighting its potential use in mobile health applications or tele dermatology platforms. V. Ravi et al. [48] presented Deepscan, a hybrid pipeline that integrates convolutional neural networks with Vision Transformers (ViTs) alongside a segmentation module. The approach exhibited remarkable classification accuracy on the proprietary DermVisD dataset. The hybrid approach employs convolutional neural networks to derive local spatial features and vision transformers to interpret global context, with segmentation improving the localisation of lesions leading to enhanced performance on small, irregular diseases.

I. Pacal et al. [49] formulated this concept by introducing a CNN–ViT hybrid architecture specifically designed for the early detection of melanoma. The model combines ResNet-50 for feature encoding with ViT blocks for attention mapping, resulting in a balanced integration of accuracy and contextual insight. Evaluations performed on the ISIC-2018 and HAM10000 datasets revealed improved results when compared to standalone ViT or CNN models, particularly in reducing false negatives, a crucial factor in melanoma diagnosis. While these advanced models show promise, the current body of work, especially from 2023 to 2025, reveals various shortcomings. Limited assistance for various categories: A significant number of studies focus on binary classification (melanoma versus non-melanoma), while real-world clinical scenarios often require the ability to distinguish between multiple disease types. Absence of segmentation integration: Limited models integrate lesion segmentation, such as UNet or UNet++, which is crucial for delineating lesion regions and enhancing classification precision. The concept of explainability continues to develop. Although some studies apply XAI tools like Grad-CAM, the incorporation of explainable segmentation and classification pipelines is still not thoroughly investigated. The proposed model effectively tackles these gaps by combining CNN, VGG-16 and ResNet technique with DA and UNet++ segmentation technique to classify multi-class lesions. Implementing methods like data augmentation and dimensionality reduction to tackle challenges related to class imbalance and noise. Concentrating on seven clinically important types of skin lesions, improving alignment with the practical requirements of dermatology. Aiming to provide clear and prompt predictions to ease the diagnostic burden faced by dermatologists.

3. Proposed strategy and baseline methodology

The strategy and methodology play a very important role in the field of research work. This research work used different methods to develop a robust model for the classification of skin lesions as explained below:

3.1. Convolutional Neural Network (CNN)

A Convolutional Neural Network (CNN) [14] is a deep learning architecture that autonomously extracts hierarchical features from input images by producing feature maps via convolutional filters. Convolutional Neural Networks (CNNs) use a feed-forward design, generally consisting of convolutional layers, pooling layers, and fully connected layers. These interrelated layers facilitate the effective acquisition of spatial hierarchies within visual data. Fig. 1 depicts a general model of a CNN architecture, designed to elucidate the standard components involved specifically, convolution, pooling, and fully connected layers arranged sequentially for feature extraction and classification. It does not represent the precise architecture implemented in Case 1. The CNN architecture deployed in Case 1 has been explained in the Methodology section, which describes specific characteristics of the model, including the number of layers, filter sizes, and activation functions.

3.2. Transfer learning

Transfer learning [23,24] is a deep learning technique in which a model that has been pre-trained on one task, often using a large-scale dataset such as ImageNet, is adapted for a different yet related task. This process involves transferring knowledge gained from addressing the source problem to enhance learning efficiency and performance on the target problem. Transfer learning effectively shortens training duration, alleviates overfitting concerns particularly with limited datasets, and improves model precision. It is crucial to recognise that data normalisation constitutes a separate preprocessing step, implemented before the training phase. This process assures identical input distributions through the scaling or standardization of image pixel values, and it is crucial irrespective of the model architecture or training approach employed. This investigation employs transfer learning with two pre-trained models, VGG16 and ResNet50, to classify skin cancer images, employing their feature extraction capabilities to improve diagnostic precision.

3.2.1. VGG architecture

The VGG architecture is a commonly employed Convolutional Neural Network (CNN) model for image classification tasks [15]. VGG16 features a profound architecture characterized by numerous convolutional layers with small receptive fields (3×3), interspersed with max pooling layers (2×2), and culminating in fully connected layers. This architecture allows effective feature extraction while maintaining computational efficiency. Fig. 2(a) provides a clear overview of the main elements of the VGG16 architecture. Concentrating on the ongoing application of convolutional, pooling, and fully connected layers. This investigation takes the VGG16

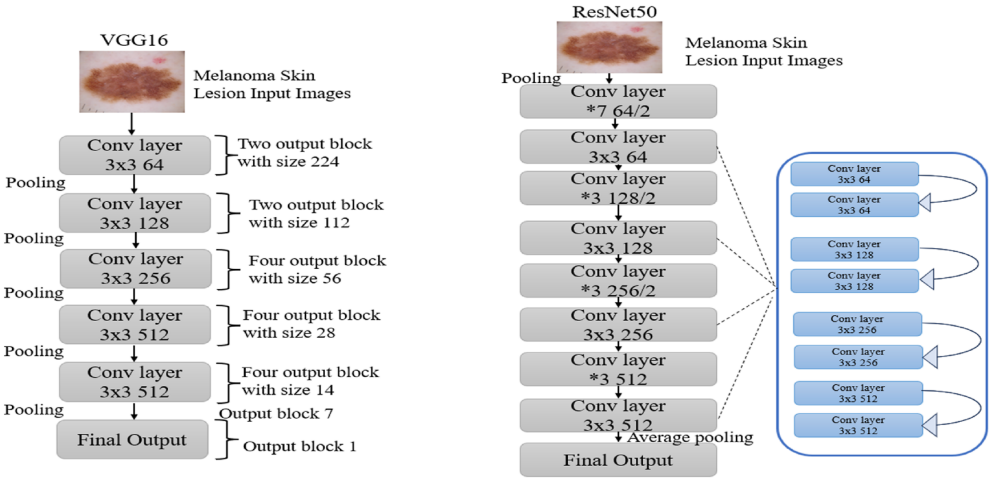

Fig. 2. Overview of the (a) VGG16 and (b) Resnet50 architecture.

model for the classification of skin images, highlighting its strong performance in visual recognition tasks. The models exhibit differences in the number of layers filter configurations and structural characteristics across the various blocks [16].

3.2.2. *ResNet architecture*

To enhance performance during training on a large dataset using a CNN architecture, it is often necessary to supplement the model with additional layers. However, increasing the depth of the network can lead to the vanishing or exploding gradient problem, which hampers the learning process and reduces the overall accuracy rate [17]. To overcome this challenge, the Residual Network (ResNet) architecture introduces residual connections, also known as shortcut connections. Rather than learning a direct underlying mapping, ResNet enables the network to learn residual functions, the difference between the input and the output of a layer. These shortcut paths allow gradients to bypass certain layers and flow unimpeded through the network, thus preserving gradient strength during backpropagation and mitigating the vanishing gradient problem. Fig. 2(b) illustrates the introduction of the ResNet design, which effectively overcomes the gradient issue and contributes to a higher level of accuracy. By enabling the construction of much deeper networks while preserving training stability, ResNet enhances model optimization and performance [18].

3.3. *Adam optimizer*

The Adam optimizer is a widely used optimization algorithm in deep learning due to its ability to adaptively adjust learning rates for individual parameters, which helps accelerate convergence and improves training stability. While Adam enhances model performance by efficiently minimizing loss, it does not inherently detect overfitting. Overfitting is typically monitored through validation metrics such as loss or accuracy on unseen data. Additionally, Adam operates on gradients, not "normalized parameters (0–1)" as previously stated. In our U-Net++ architecture equipped with pixel-level and spatial multi-scale attention modules Adam plays a crucial role in refining feature maps during backpropagation and weight updates [39]. The model further integrates image segmentation and edge perception networks within a rough-to-fine encoding-decoding structure. The edge perception network is trained to learn residuals between predicted and true edges using a loss function, thereby improving the precision of segmentation.

3.4. *Data augmentation*

To accurately categorize images as cancerous or not, image classification is frequently utilized in the medical industry, which contains a huge amount of data to evaluate and train the model. Due to the large volume of image datasets, all images are not consistent and vary in many ways, including size, orientation, and color. The results of the trained model produce inaccurate results due to the inconsistency of datasets [19]. The augmentation method uses duplicates of the original data to address data discrepancies and recognize targets of varying sizes and shapes [22]. This method effectively prevents the training model from overfitting. Stratification serves as a supplementary method applied alongside the data. Augmentation methods

aimed at addressing skewed samples, in contrast to stratification techniques, prevent additional data splitting post-initial randomisation. They establish conditions for subsequent data division into training and testing groups that are both random and devoid of bias in the data splitting process. The data augmentation technique described by C.N. Vasconcelos et al. [21] allows for the images in the melanoma dataset to be flipped, rotated, and cropped to improve accuracy.

In this study, we addressed this imbalance using a combination of strategies beyond standard data augmentation. First, we employed class-balanced sampling during training to ensure each mini-batch contains a fair representation from underrepresented classes. This helped mitigate bias toward majority classes during gradient updates. Second, a weighted categorical cross-entropy loss function was applied, where higher weights were assigned to minority classes based on the inverse frequency of class distribution. This approach penalized the misclassification of underrepresented classes more heavily, improving model generalization across all categories. These techniques, when combined with extensive data augmentation, provided a more equitable learning process and reduced the risk of overfitting to dominant classes. The effectiveness of these balancing methods is reflected for improving the performance of models.

3.5. UNet++ technqiue

The UNet++ technique is a deep learning method that combines datasets and segmentation techniques to improve image segmentation accuracy. UNet++ extends the standard U-Net architecture by introducing nested and densely connected skip pathways between encoder and decoder blocks. These enhanced skip connections reduce the semantic gap between high-resolution encoder features and low-resolution decoder features, leading to more precise segmentation. It uses a two-section approach: the contracting path employs a conventional CNN for feature extraction, followed by convolutions. However, unlike the standard U-Net, UNet++ adds intermediate convolutional layers within the skip connections, enabling better feature refinement at each stage of the decoder. ReLU activates the convolutions after up sampling, reducing, and linking the feature map. Unlike the original U-Net, which required cropping to align feature map dimensions due to padding mismatches, UNet++ employs nested skip pathways that reduce the reliance on cropping and preserve contextual information across scales, all while maintaining the U-shaped structure. The skin detection and segmentation phase uses minimal training samples and generates detailed segmentation feature maps [40]. Data augmentation helps improve model performance, particularly in medical imaging tasks like skin cancer diagnosis. UNet++ gains greatly from enriched data, which prevents overfitting and boosts dataset diversity. This addition improves generalization to data that hasn't been seen before and makes it possible to extract finer features, enhance edge detection, close semantic gaps, and handle difficult variations better. The UNet++ technique improves image segmentation accuracy, especially in medical imaging where lesions may change in form, size, and texture. UNet++ is employed in this study as a preprocessing module to enhance classification performance by accurately localizing the lesion region in dermoscopic images. Specifically, UNet++ generates a binary segmentation mask that delineates the lesion from the surrounding skin. This mask is then used in the following way: the original image is element-wise multiplied with the UNet++ mask, effectively filtering out background noise and preserving only the lesion area. The resulting masked image serves as the input to the VGG16 and ResNet50 classification models, replacing the original RGB image. This masked image retains the same dimensions as the original input, and no additional input channel is introduced. Instead, the segmentation step enhances the discriminative focus of the classifier by eliminating irrelevant features outside the lesion boundary. Through this integration strategy, the UNet++ segmentation model acts as an attention mechanism, directing the classifier toward the most diagnostically relevant region of interest (ROI). This preprocessing is applied only in Case 3, allowing for a comparative evaluation of model performance with and without segmentation.

4. Data source

This research work used three types of datasets that were collected from the Kaggle open-source repository. The first dataset is the HAM10000 dataset which contains a total of 10015 Images of skin cancer disease. Images of the HAM10000 dataset were divided into seven classes and included seven different types of skin lesions like melanoma (mel), melanocytic nevi (nv), actinic keratoses, and intraepithelial carcinoma/Bowen disease (apiece), basal cell carcinoma (bcc), benign lesions of the keratosis type (solar lentigines/seborrheic keratoses and lichen-planus like keratosis,

bkl), dermatofibroma (pdf), melanoma (mel), melanocytic nevi (nv), and vascular lesions (angiomas, angiokeratomas, pyogenic granulomas and hemorrhages, vasc), where melanoma (mel), basal cell carcinoma (bcc) is the type of skin cancer, and the remaining are non-skin cancer [36]. Table 1 shows the HAM10000 skin Lesions dataset with multiple classes.

The second dataset is the Malignant vs. Benign ISIC Archive dataset, which includes a total 3297 images where 1497 images belong to malignant (cancer) and 1800 images belong to benign (non-cancer) [35]. Table 2 shows that Malignant vs. Benign ISIC Archive dataset with binary classification.

The third dataset is called the Melanoma Skin Cancer dataset which includes a total of 10605 images l, having 5105 images of malignant (cancer) and 5500 images of benign (non-cancer) samples [11]. Table 3 shows that Melanoma Skin Cancer dataset with binary classification.

5. Implementation details

The experimental analysis for this study was conducted using Python (version 3.8) as the core programming language. We employed several prominent deep learning libraries, including TensorFlow (open-source) and Keras, for model development and training. Scikit-learn, NumPy, Matplotlib, and Pandas were used for preprocessing, statistical analysis, and visualization tasks. Python's extensive ecosystem of libraries, such as Pandas and Matplotlib, enabled efficient data manipulation, statistical analysis, and generation of meaningful visual diagrams. Model training and evaluation were performed using Jupiter Notebook locally and Google Collaboratory (Collab) as an integrated development environment. Collab's cloud-based GPU support provided an accessible and scalable solution for training deep learning models, bypassing the need for high-end local hardware. The models were executed on Collab's virtual GPU (approximately 2 GB VRAM), enabling faster computation while the datasets were stored and accessed via Google Drive. The local system used for initial prototyping was an AMD Ryzen 5 notebook with 8 GB RAM and no dedicated GPU, but the final training and validation were conducted on Google Collab, ensuring resource efficiency and compatibility with larger datasets. TensorFlow's machine learning operations were executed via its dataflow model, facilitating optimized training through efficient value computation and backpropagation. To address overfitting and improve generalization, the implementation included data augmentation, dropout layers, and early stopping. Furthermore, cross-validation was adopted to validate the robustness and consistency of the models across multiple data partitions [54].

Table 1. HAM10000 skin Lesions dataset with multiclass classification.

MEL	NV	BCC	AKIEC	BKL	DF	VASC
113	6705	514	327	1099	115	142

Table 2. Malignant vs. Benign ISIC Archive dataset with binary classification.

Category of Sample	Number of samples
Malignant	1497
Benign	1800

Table 3. Melanoma Skin Cancer dataset with binary classification.

Category of Sample	Number of samples
Malignant	5105
Benign	5500

6. Experimentation setup

Model training in deep learning presents several challenges that can impact the overall performance and generalization of the model. One of the common issues is overfitting, where the model achieves high accuracy on training data but performs poorly on validation data, indicating it has memorized the training set rather than learning generalizable features. Conversely, underfitting occurs when the model fails to capture the underlying patterns in the data, resulting in low accuracy on both training and validation sets. These limitations are particularly significant in medical image analysis tasks, where annotated data is limited and variations in lesion appearance are complex. To address these challenges, several strategies were adopted in this study. Data augmentation techniques such as rotation, flipping, scaling, and color jittering were used to artificially increase dataset diversity, thereby reducing overfitting by exposing the model to various forms of the same lesion [50]. Early stopping was employed during training to monitor validation loss and halt training once the performance on the validation set began to deteriorate, helping to prevent overfitting [51]. Dropout layers were incorporated within the deep learning architecture, particularly in the fully connected layers, to randomly deactivate a fraction of neurons during training. This forces the network to learn redundant representations and discourages over-reliance on

specific features [52]. K-fold cross-validation was applied to ensure the model's robustness across different data splits. This approach divides the dataset into k subsets, trains the model on k−1 subsets, and validates it on the remaining one, rotating this process. It provides a more comprehensive evaluation and reduces the risk of both overfitting and underfitting [53]. By integrating these techniques, the model training process was optimized for stability, generalization, and improved diagnostic accuracy, particularly in tasks involving skin cancer classification and segmentation. The experimental setup in the initial version of this study was described too vaguely to support reproducibility. In this manuscript, we now clearly specify all relevant hyperparameters, computational environments, and training configurations. The models were trained using the Adam optimizer with an initial learning rate of 0.0001, and a learning rate scheduler was employed to reduce the learning rate by a factor of 0.1 every 60 epochs without improvement in validation loss. The batch size was set to 64, and training was conducted for 100 epochs with early stopping patience of 60 epochs to avoid overfitting. Regarding data augmentation, we applied a range of transformations using the Keras Image Data Generator, including Random rotation within the range of ±20°, Horizontal and vertical flipping, width and height shift up to 10 %, Zoom range from 0.8 to 1.2, Shear intensity of 0.2, Brightness range of [0.8, 1.2]. These augmentations were applied only to the training set, ensuring the test set remained untouched for unbiased evaluation.

6.1. Observations and decisions based on experimental results

The experiment is divided into three distinct cases: case 1, case 2, and case 3. The following Fig. 3 shows the flow of the classification process of skin and non-skin cancer diseases with three different cases. In the first case, skin and non-skin cancer diseases are classified with all features. In the second case, the skin and non-skin cancer diseases are classified with data augmentation technique. In the third case, skin and non-skin cancer diseases are classified with the U-net+ model and data augmentation technique.

Case 1. This case involves the classification of skin cancer with the original dataset. In Case 1, the proposed flow architecture is divided into three sections: preprocessing of the dataset, training, and testing of the dataset, and evaluation and comparison of models. We have resized and cropped images of the skin cancer dataset which is an essential and initial action of preprocessing while preparing images. It is then followed by adjusting the pixel values. A commonly used method for normalizing is to adjust the pixel values so that they fall within the range of 0−1. This is achieved by standardizing all pixels by the maximum pixel value (often 255 for RGB images), after the application of noise reduction and sharpening filters to the images. Feature extraction is also very essential to improve the performance of the model. In this process, the Gaussian filter method feature extraction is applied after the normalization of images. The preprocessed dataset is divided into training and testing partition, where training and testing datasets are used to train and test the deep learning-based models respectively, and finally evaluates and compares the performance of models for classification of skin and non-skin cancer images.

Our research conducted a comprehensive and methodical analysis across three datasets, as indicated in Table 4, which presents the details of the classification technique. In this work, the training and testing samples of three different skin datasets

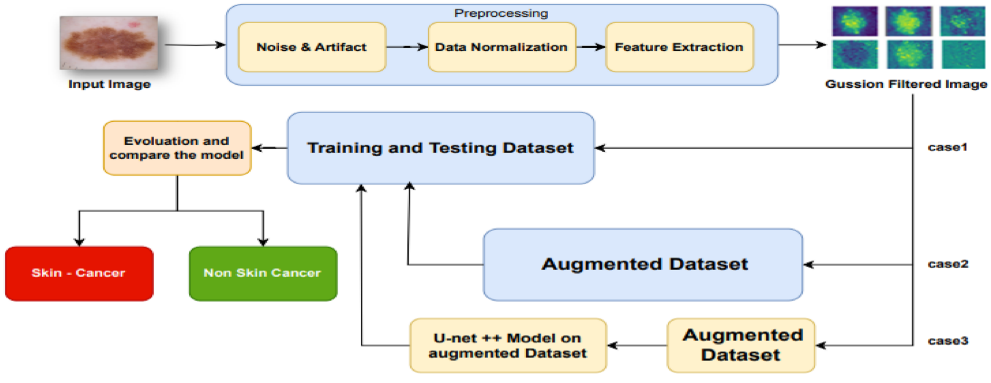

Fig. 3. Flow of the proposed architecture for classification of skin cancer disease.

Table 4. Performance comparison of transfer learning with different datasets.

Model's	Ham1000 Dataset		Melanoma Skin Cancer Dataset		Malignant vs. Benign ISIC Archive dataset	
	Training	Testing	Training	Testing	Training	Testing
CNN	97.63 %	87.29 %	94.25 %	86.24 %	96.78 %	88.28 %
ResNet 50	96.82%	93.29 %	**95.65 %**	**93.97 %**	94.97 %	91.20 %
VGG 16	**98.26 %**	**96.07 %**	95.24 %	92.87 %	**94.61 %**	**94.55 %**

Bold value indicates the highest performance measures.

were applied to three deep learning-based models: CNN, ResNet50, and VGG16. We used the Adam optimizer to evaluate and optimize performance. The key evaluation metric used for model comparison was accuracy measured on both training and testing datasets. Testing accuracy served as the primary metric for final model selection, as it reflects generalization performance. The results show that the VGG16 model achieved the best overall performance, particularly on the HAM10000 dataset, with 98.26 % training accuracy and 96.07 % testing accuracy. It also performed well on the ISIC Archive dataset (Malignant vs. Benign), with 94.61 % training accuracy and 94.55 % testing accuracy, demonstrating robustness in binary classification tasks. For the Melanoma Skin Cancer dataset, ResNet50 achieved the highest accuracy, recording 95.65 % training accuracy and 93.97 % testing accuracy. This suggests that ResNet50 may be more effective for datasets that focus specifically on melanoma detection, likely due to its residual learning capabilities which help in deeper feature extraction and reducing vanishing gradient issues. CNN showed the lowest performance across all datasets, with significantly lower testing accuracies (e.g., 87.29 % on HAM10000 and 86.24 % on the Melanoma dataset), indicating possible overfitting or insufficient capacity for deeper feature representation. Category-wise, while the metrics are reported at the dataset level, inference can be made based on dataset characteristics: The ISIC Archive dataset, which is binary (malignant vs. benign), saw the best performance by VGG16, suggesting it handled both benign and malignant classes effectively. On the Melanoma Skin Cancer dataset, ResNet50's superior performance implies better malignant lesion detection. CNN struggled to generalize, especially for the more complex or imbalanced datasets. Based on these results, decisions can be drawn as follows: VGG16 is the best choice for multi-class and binary classification problems, due to its consistently high testing accuracy. ResNet50 is the recommended model for datasets focusing on melanoma or malignant lesion detection. CNN should not be considered for clinical deployment without significant architecture enhancement or preprocessing adjustments, given its lower generalization. These decisions are critical for real-world deployment in medical image classification, where high accuracy, especially in detecting malignant lesions, is vital for early diagnosis and patient outcomes.

Case 2. Case 2 involves the classification of skin cancer using an enhanced pipeline that incorporates data augmentation. The Case 2 flow architecture adds a key preprocessing step as data augmentation while the rest of the model training pipeline remains similar to Case 1. By employing this technique, the focus shifted to improving model performance using novel transformations to the dataset. To enhance the training data diversity and balance, we applied random rotation, flipping, and cropping techniques to the original images. This not only increased the volume of the training dataset but also allowed the models to generalize better by seeing different perspectives of the same lesions. This adjustment in data distribution proved beneficial for the downstream classification process. The augmented datasets were then used to train three deep learning models using transfer learning like CNN-DA, ResNet50-DA, and VGG16-DA, where "DA" denotes the integration of data augmentation with deep learning and transfer learning technqiues. All models were optimized using the Adam optimizer and evaluated across three different skin cancer datasets like HAM10000, Melanoma Skin Cancer, and ISIC Archive (Malignant vs. Benign). The performance results are displayed in Table 5. The VGG16-DA model achieved the highest overall performance, registering 98.89 % training accuracy and 98.63 % testing accuracy on the HAM10000 dataset, and 98.02 % training and 97.43 % testing accuracy on the ISIC dataset. These results underscore the model's capability to handle both multi-class and binary classifications effectively, even when trained on augmented data. For the Melanoma Skin Cancer dataset, ResNet50-DA emerged as the best-performing model, obtaining 98.45% training accuracy and 96.77 % testing

Table 5. Performanc e comparison of transfer learning models using data augmentation with different datasets.

Model's	Ham1000 Dataset		Melanoma Skin Cancer Dataset		Malignant vs. Benign ISIC Archive dataset	
	Training	Testing	Training	Testing	Training	Testing
CNN-DA	97.83 %	88.39 %	94.75 %	87.48 %	97.58 %	89.88 %
ResNet50-DA	98.64 %	97.86 %	**98.45 %**	**96.77 %**	97.77 %	94.60 %
VGG16-DA	**98.89 %**	**98.63 %**	97.84 %	96.67 %	**98.02 %**	**97.43 %**

Bold value indicates the highest performance measures.

accuracy. The use of residual connections likely helped in better learning of deeper, non-linear patterns, particularly useful in malignant lesion detection. On the other hand, CNN-DA showed modest improvements compared to its non-augmented counterpart (CNN from Case 1), but still underperformed compared to ResNet50-DA and VGG16-DA. This suggests that despite the benefits of augmentation, model depth and architecture complexity remain critical for high-accuracy lesion classification. Data augmentation led to significant improvements in generalization across all models and datasets. VGG16-DA consistently delivered the best results overall, making it highly suitable for clinical deployment. ResNet50-DA is ideal for malignant melanoma detection, due to its strength in learning complex features. CNN-DA improved modestly, but remains less competitive due to its shallower architecture.

Case 3. Case 3 involves the classification of skin cancer through a hybrid approach that integrates both data augmentation and the UNet++ segmentation technique. The Case 3 flow architecture introduces a key additional step UNet++ segmentation building upon the Case 2 pipeline. This enhanced preprocessing pipeline seeks to improve model accuracy and robustness by producing more refined input for the learning phase. To further boost the model's performance, we combined data augmentation with UNet++ segmentation, enabling the model to extract more fine-grained lesion features. Data augmentation, already proven effective in Case 2, remains valuable in increasing model generalization, especially in medical imaging tasks like skin lesion classification. UNet++ enhances the model's ability to perform pixel-level segmentation with high precision, even from limited training samples, which is particularly crucial in medical applications. In our segmentation phase, UNet++ generates highly detailed segmentation feature maps, which help the classification models (CNN, ResNet50, VGG16) focus more accurately on lesion-specific regions. In this case, we proposed three different hybrid models:

- **CNN-DAUNPP**: CNN + Data Augmentation + UNet++
- **ResNet50-DAUNPP**: ResNet50 + Data Augmentation + UNet++
- **VGG16-DAUNPP**: VGG16 + Data Augmentation + UNet++

All three models were evaluated on three benchmark skin cancer datasets—HAM10000, Melanoma Skin Cancer, and ISIC Archive (Malignant vs. Benign), and optimized using the Adam optimizer. The performance outcomes are presented in Table 6. VGG16-DAUNPP is the best overall model, showing consistently superior training and testing accuracies, particularly 98.83 % testing accuracy on HAM10000 dataset and 98.63 % on ISIC dataset (Benign vs Malignant). ResNet50-DAUNPP delivered the best results for the Melanoma Skin Cancer dataset, especially in malignant lesion classification, achieving 98.87 % testing accuracy, making it ideal for early and accurate detection of life-threatening cancers. CNN-DAUNPP, despite improvements over Case 1, did not benefit significantly from the UNet++ segmentation, indicating that simpler CNN architectures may not fully utilize the rich feature maps provided by UNet++. All performance evaluations were primarily based on the accuracy metric, as shown in Table 6. However, for

Table 6. Performance comparison of transfer learning models using UNet++ segmentation technique on augmented data with different datasets.

Model's	Ham1000 Dataset		Melanoma Skin Cancer Dataset		Malignant vs. Benign ISIC Archive dataset	
	Training	Testing	Training	Testing	Training	Testing
CNN-DAUNPP	97.83 %	88.39 %	94.75 %	87.48 %	97.58 %	89.88 %
ResNet50-DAUNPP	98.64 %	97.86 %	**98.45 %**	**98.87 %**	97.77 %	94.60 %
VGG16-DAUNPP	**98.89 %**	**98.83 %**	97.84 %	96.67 %	**98.02 %**	**98.63 %**

Bold value indicates the highest performance measures.

clinical diagnosis, F1-score and recall will be highlighted in future work due to their importance in handling imbalanced data scenarios. The UNet++ segmentation step significantly improved localization and boundary detection, which in turn enhanced classification accuracy especially critical in detecting subtle malignancy patterns. Decisions informed by these Observations, VGG16-DAUNPP is highly recommended for deployment in multi-class and binary classification tasks due to its robustness and adaptability across datasets. ResNet50-DAUNPP should be used where the detection of malignant lesions (e.g., melanoma) is prioritized, given its depth and feature refinement. CNN-DAUNPP is not recommended for clinical use, unless model complexity is strictly limited by computational constraints.

In Fig. 4, the validation loss curve appears to begin at a lower value than the training loss curve during the initial epochs, which might seem counterintuitive. This behavior can occur due to the presence of regularization techniques (such as dropout), which are active only during training and not during validation. During the early stages of training, the model may perform better on the relatively simpler validation data (which is not augmented), while struggling with the augmented training samples. As training progresses, the model adapts and this discrepancy typically resolves. This behavior is well-documented in deep learning literature.

Figs. 4 and 5 show the accuracy and model loss of the proposed CNN-DAUNPP with Ham1000 and Malignant vs. Benign ISIC Archive dataset respectively. Fig. 6 shows the accuracy and model loss of

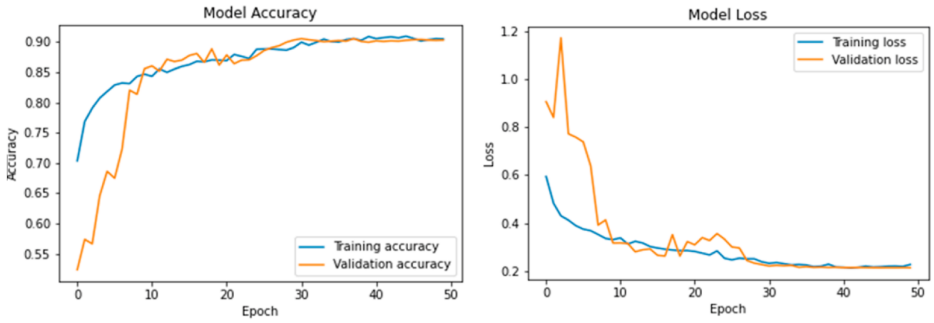

Fig. 4. Accuracy and model loss of VGG16-DAUNPP with Ham1000 dataset.

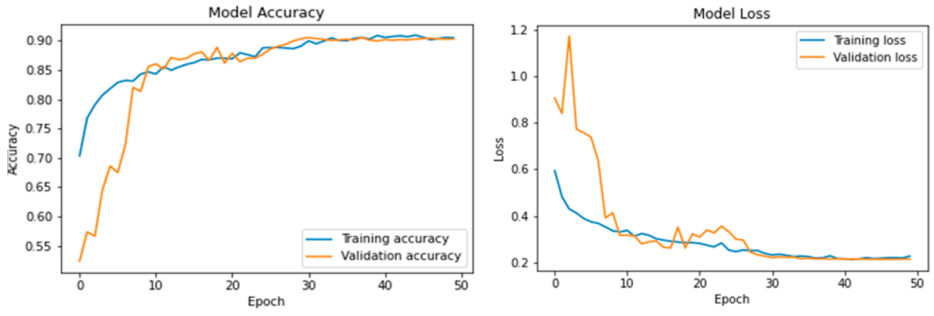

Fig. 5. Accuracy and model loss of VGG16-DAUNPP with Malignant vs. Benign ISIC archive dataset.

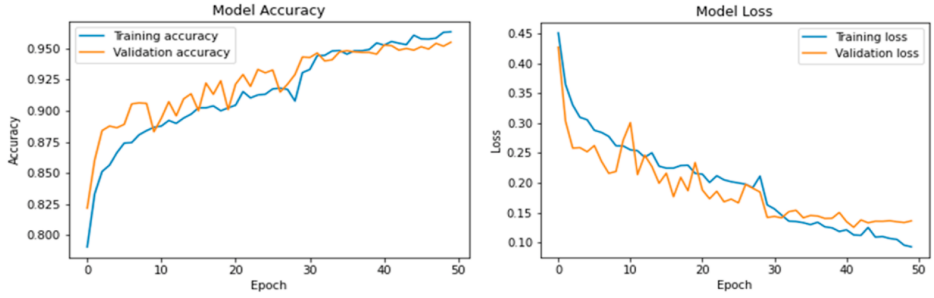

Fig. 6. Accuracy and model loss of ResNet50-DAUNPP with Melanoma skin cancer dataset.

Table 7. A comparative analysis of relative works with the proposed model.

Author Name and Year	Dataset	Deep Learning models	Accuracy (%)
Gamage, L. et al. (2024) [55]	ISIC2019	EfficientNet-B4 with DenseNet Feature Fusion	92.68 %
Gouda, W. et al. (2022) [28]	ISIC2019	CNN, ResNet-50	83.1 % with CNN,83.6 % with ResNet-50
Srinivasu, P. N. et al. (2021) [8]	HAM10000	Mobile Net with LSTM	85 %
Yao, P. et al. (2021) [29]	HAM10000	RegNetY-3.2GF	85.8 %
Rahman, Z. et al. (2021) [30]	HAM10000	ResNetXt, SeResNeXt, DenseNet, Xception, ResNet	88 %, 89 %, 91 %, 88 % & 84 % respectively
Ameri, A. (2020) [31]	HAM10000	AlexNet	84 %
Polat, K., & Koc, K. O. (2020) [32]	HAM10000	CNN, ResNet	77 % with CNN, 78 % with ResNet
Sae-Lim, W. et al. (2019) [33]	HAM10000	MobileNet	83.9 %
Li, X. et al. (2019) [34]	ISCI2019	AlexNet,Vgg16, ResNet18, ResNet101	80.31, 84.16, 85.65 & 85.54 % respective
Proposed Model	ISIC2019	VGG16-DAUNPP	98.63 %
	HAM10000,	VGG16-DAUNPP	98.83 %
	Melanoma Skin Cancer Dataset	ResNet50-DAUNPP	98.87 %

the proposed ResNet50-DAUNPPmodel with the Melanoma Skin Cancer dataset.

7. Result analysis and comparison

This section explores the comparative analysis of the proposed model with other existing models work done by different researchers with HAM10000, Melanoma Skin Cancer, and Malignant vs. Benign ISIC Archive dataset in terms of accuracy. Consequently, it's reasonable to assume that the data augmentation and UNet++ segmentation techniques help to enhance the performance of the model. Table 7 shows that comparative analysis of relative works with our proposed model where our proposed model achieved better accuracy as compared to others.

8. Conclusion

In this study, we conduct a thorough empirical assessment of deep learning frameworks for the categorisation of skin cancer using dermoscopy-based approaches. Deep learning models may be able to help medical practitioners discover skin lesions, which is an important first step in the process of caring for patients. Augmentation and UNet++ segmentation approaches are particularly significant for increasing the performance of deep learning-based models. This research work develops a robust deep learning-based model using CNN, ResNet50, and VGG16 with data augmentation and UNet++ segmentation techniques. The proposed models are compared with and without data augmentation and UNet++ segmentation technique where our proposed deep learning-based model achieved better accuracy with data augmentation and UNet++ segmentation techniques. The proposed model is compared in terms of accuracy where the VGG16-DAUNPP model achieves better accuracy with Ham1000 and with Malignant vs. Benign ISIC Archive dataset in case of both with and without data augmentation and UNet++ segmentation technique, but ResNet50 achieves better accuracy with Melanoma Skin Cancer dataset in case of both with and without data augmentation and UNet++ segmentation technique. We have also compared our proposed model with previously developed models by different authors in which our proposed model achieved better accuracy as compared to others. In conclusion, our proposed model is not only effective and accurate for skin cancer classification but also shows strong potential for real-world clinical deployment, addressing the original motivation of cost-effective, objective, and scalable diagnosis.

9. Future scope and clinical deployment

In future work, we aim to explore additional optimization algorithms, ensemble fusion strategies, and image enhancement techniques to further improve classification accuracy as well as validate other performance measures like sensitivity, specificity, precision and F-score across diverse datasets and skin types. Moreover, to bridge the gap between research and clinical practice, we will work towards validating the model's predictions through collaboration with

dermatologists and other healthcare professionals. Inspired by studies such as L. Gamage et al. [55], which discuss real-world implementation challenges, our goal is to deploy the proposed framework as a practical computer-aided diagnostic tool in dermatology clinics. Such deployment will involve rigorous clinical trials, integration with electronic health record systems, and user-interface development for clinicians. By addressing model interpretability, reliability, and regulatory compliance, our future work will support the translation of AI-based skin cancer classification systems from academic prototypes to usable medical applications. Furthermore, to evaluate the clinical applicability of the proposed model, collaboration with medical experts such as dermatologists is essential. Although this study primarily focuses on algorithmic performance across benchmark datasets, future validation will include expert annotation, double-blinded testing, and histopathological ground truth comparison. Such collaboration would help verify the model's diagnostic decisions in real-world scenarios and enhance its trust and usability in clinical workflows. Additionally, clinician feedback will guide necessary improvements in usability, interpretability, and clinical safety of the deployed system. In future we will also plan to incorporate and train advanced CNN and Vision Transformer (ViT) models such as ConvNeXt, Swin Transformer, DeiT, ViT, and EfficientNetV2 on the same datasets and experimental setups. This approach will increase accessibility, especially in remote and underserved areas, aligning with the broader goal of democratizing healthcare through AI.

Ethical statement

This article does not contain any studies with human participants or animals performed by any of the authors.

Funding

No funds.

Conflicts of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgements

The authors acknowledges to my colleague Dr. Santosh Majhi, Associate Professor Guru Ghasidas Vishwavidyalaya Bilaspur, India for suggesting relevant general to submit our article.

References

[1] M.A. Al-masni, M.A. Al-antari, M.T. Choi, S.M. Han, T.S. Kim, Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks, Comput. Methods Programs Biomed. 162 (2018) 221–231, https://doi.org/10.1016/j.cmpb.2018.05.027. Get rights and content.

[2] C. Ananth, M.J. Therese, A survey on melanoma: skin cancer through computerized diagnosis, Int. J. Adv. Res. Innov. Discov. Eng. Appl. (IJARIDEA) 5 (2020) 9–18, https://doi.org/10.2139/ssrn.3551811.

[3] K.M. Hosny, M.A. Kassem, M.M. Foaud, Classification of skin lesions using transfer learning and augmentation with Alex-Net, PLoS One 14 (2019) 1–17, https://doi.org/10.1371/journal.pone.0217293.

[4] D. Gangwani, P. Gangwani, Applications of machine learning and artificial intelligence in intelligent transportation system: a review, in: Applications of Artificial Intelligence and Machine Learning: Select Proceedings of ICAAAIML, 778, 2021, pp. 203–216, https://doi.org/10.1007/978-981-16-3067-5_16.

[5] K. Matsunaga, A. Hamada, A. Minagawa, H. Koga, Image classification of melanoma, nevus and seborrheic keratosis by deep neural network ensemble, arXiv preprint arXiv: 1703.03108 (2017) 2–9, https://doi.org/10.48550/arXiv.1703.03108.

[6] M.A. Milton, Automated skin lesion classification using ensemble of deep neural networks in ISIC 2018: skin lesion analysis towards melanoma detection challenge, arXiv preprint arXiv:1901.10802 (2019) 4–15, https://doi.org/10.48550/arXiv.1901.10802.

[7] Q. Gan, L.S. Wei, T. Ji, Skin disease recognition method based on image color and M. F texture features, Comput. Math. Methods Med. 1 (2018) 123–132, https://doi.org/10.1155/2018/8145713.

[8] P.N. Srinivasu, J.G. SivaSai, M.F. Ijaz, A.K. Bhoi, W. Kim, J.J. Kang, Classification of skin disease using deep learning neural networks with MobileNet V2 and LSTM, Sensors 21 (2021) 2852–2879, https://doi.org/10.3390/s21082852.

[9] S. Malliga, G.S. Infanta, S. Sindoora, S. Yogarasi, Skin disease detection and classification using deep learning algorithms, Int. J. Adv. Sci. Technol. 29 (2020) 255–360. https://sersc.org/journals/index.php/IJAST/article/view/5587.

[10] N. Hameed, A.M. Shabut, M.A. Hossain, Multi-class skin diseases classification using deep convolutional neural network and support vector machine, in: Proc. 12th IEEE Int. Conf. Softw., Knowl. Inf. Manage. Appl. (SKIMA), 2018, pp. 1–7, https://doi.org/10.1109/SKIMA.2018.8631525.

[11] Dataset Source: "Melanoma skin cancer dataset of 10000 images," [Online] collected from Kaggle repository. [Accessed: July-07-2023].

[12] T. Sreelatha, M.V. Subramanyam, M.G. Prasad, Early detection of skin cancer using melanoma segmentation technique, J. Med. Syst. 43 (2019) 190–197, https://doi.org/10.1007/s10916-019-1334-1.

[13] M.Q. Khan, A. Hussain, S.U. Rehman, U. Khan, M. Maqsood, K. Mehmood, M.A. Khan, Classification of melanoma and nevus in digital images for diagnosis of skin cancer, IEEE Access 7 (2019) 90132–90144, https://doi.org/10.1109/ACCESS.2019.2926837.

[14] A.R. Lopez, X. Giro-i-Nieto, J. Burdick, O. Marques, Skin lesion classification from dermoscopic images using deep learning techniques, in: Proc. 13th IASTED Int. Conf. Biomed. Eng. (Biomed), 2017, pp. 49–54, https://doi.org/10.2316/P.2017.852-053.

[15] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556 1 (2015) 1–14, https://doi.org/10.48550/arXiv.1409.1556.

[16] J. Jaworek-Korjakowska, P. Kleczek, M. Gorgon, Melanoma thickness prediction based on convolutional neural network with VGG-19 model transfer learning, in: Proc. IEEE/CVF

Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), 2019, pp. 2748–2756, https://doi.org/10.1109/CVPRW.2019.00333.
[17] E. Perez, O. Reyes, S. Ventura, Convolutional neural networks for the automatic diagnosis of melanoma: an extensive experimental study, Med. Image Anal. 67 (2021) 1–13, https://doi.org/10.1016/j.media.2020.101858.
[18] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, pp. 770–778, https://doi.org/10.1109/CVPR.2016.90.
[19] D. Bisla, A. Choromanska, R.S. Berman, J.A. Stein, D. Polsky, Towards automated melanoma detection with deep learning: data purification and augmentation, in: Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), 2019, pp. 2720–2728, https://doi.org/10.1109/CVPRW.2019.00330.
[20] P.A. Diallo, Y. Ju, Accurate detection of COVID-19 using K-EfficientNet deep learning image classifier and K-COVID chest X-ray images dataset, in: Proc. 6th IEEE Int. Conf. Comput. Commun. (ICCC), 2020, pp. 1527–1531, https://doi.org/10.1109/ICCC51575.2020.9344949.
[21] C.N. Vasconcelos, B.N. Vasconcelos, Convolutional neural network committees for melanoma classification with classical and expert knowledge-based image transforms data augmentation, arXiv preprint arXiv:1702.07025 (2017) 2012–2023, https://doi.org/10.48550/arXiv.1702.07025.
[22] A. Menegola, M. Fornaciali, R. Pires, F.V. Bittencourt, S. Avila, E. Valle, Knowledge transfer for melanoma screening with deep learning, in: Proc. 14th IEEE Int. Symp. Biomed. Imaging, 2017, pp. 297–300, https://doi.org/10.1109/ISBI.2017.7950523.
[23] J.O. Emuoyibofarhe, D. Ajisafe, R.S. Babatunde, M. Christoph, Early skin cancer detection using deep convolutional neural networks on mobile smartphone, Int. J. Inf. Eng. Electron. Bus. 12 (2020) 21–27. http://www.mecspress.org/10.5815/ijieeb.2020.02.04.
[24] E. Rezende, G. Ruppert, T. Carvalho, F. Ramos, P. De Geus, Malicious software classification using transfer learning of ResNet-50 deep neural network, in: Proc. 16th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA), 2017, pp. 1011–1014, https://doi.org/10.1109/ICMLA.2017.00-19.
[25] M.A. Marchetti, N.C. Codella, S.W. Dusza, D.A. Gutman, B. Helba, A. Kalloo, N. Mishra, C. Carrera, M.E. Celebi, J.L. DeFazio, N. Jaimes, Results of the 2016 International Skin Imaging Collaboration International Symposium on Biomedical Imaging challenge: Comparison of the accuracy of computer algorithms to dermatologists for the diagnosis of melanoma from dermoscopic images, J. Amer. Acad. Dermatol. 78 (2018) 270–277, https://doi.org/10.1016/j.jaad.2017.08.016.
[26] B. He, C. Huang, G. Sharp, S. Zhou, Q. Hu, C. Fang, Y. Fan, F. Jia, Fast automatic 3D liver segmentation based on a three-level AdaBoost-guided active shape model, Med. Phys. 43 (2016) 2421–2434, https://doi.org/10.1118/1.4946817.
[27] L.F. Li, X. Wang, W.J. Hu, N.N. Xiong, Y.X. Du, B.S. Li, Deep learning in skin disease image recognition: a review, IEEE Access 30 (2020) 208264–208280, https://doi.org/10.1109/ACCESS.2020.3037258.
[28] W. Gouda, M. Almurafeh, M. Humayun, N.Z. Jhanjhi, Detection of COVID-19 based on chest X-rays using deep learning, Healthcare 10 (2022) 343–352, https://doi.org/10.3390/healthcare10020343.
[29] P. Yao, S. Shen, M. Xu, P. Liu, F. Zhang, J. Xing, P. Shao, B. Kaffenberger, R.X. Xu, Single model deep learning on imbalanced small datasets for skin lesion classification, IEEE Trans. Med. Imaging. 41 (2021) 1242–1254, https://doi.org/10.1109/TMI.2021.3136682.
[30] Z. Rahman, A.M. Ami, A transfer learning-based approach for skin lesion classification from imbalanced data, in: Proc. 11th Int. Conf. Electr. Comput. Eng. (ICECE), 2020, pp. 65–68, https://doi.org/10.1109/ICECE51571.2020.9393155.
[31] A. Ameri, A deep learning approach to skin cancer detection in dermoscopy images, J. Biomed. Phys. Eng. 10 (2020) 801–806, https://doi.org/10.31661/jbpe.v0i0.2004-1107.
[32] K. Polat, K.O. Koc, Detection of skin diseases from dermoscopy image using the combination of convolutional neural network and one-versus-all, J. Artif. Intell. Syst. 2 (2020) 80–97, https://doi.org/10.33969/AIS.2020.21006.
[33] W. Sae-Lim, W. Wettayaprasit, P. Aiyarak, Convolutional neural networks using MobileNet for skin lesion classification, in: Proc. 16th Int. Joint Conf. Comput. Sci. Softw. Eng. (JCSSE), 2019, pp. 242–247, https://doi.org/10.1109/JCSSE.2019.88641555.
[34] X. Li, J. Wu, E.Z. Chen, H. Jiang, Deep learning towards finding skin lesion biomarkers, in: Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), 2019, pp. 2797–2800, https://doi.org/10.1109/EMBC.2019.885733.
[35] Dataset Skin-cancer-malignant-vs-benign collected from Kaggle repository (Browsing date:23-02-2024).
[36] Dataset Skin-cancer-mnist-ham10000 is collected from Kaggle repository (Browsing date:23-02-2024).
[37] K. Sathishkumar, J. Sankarapillai, S. Santhappan, A. Mathew, R.A. Nair, N. Gangane, P. Mathur, Geographic disparities in oral cancer survival from 10 population-based cancer registries in India, JAMA Netw. Open 8 (2025) 123–130, https://doi.org/10.1001/jamanetworkopen.2025.3910.
[38] K. Sathishkumar, M. Chaturvedi, P. Das, S. Stephen, P. Mathur, Cancer incidence estimates for 2022 and projection for 2025: result from National Cancer Registry Programme, India, Indian J. Med. Res. 156 (2022) 598–607, https://doi.org/10.4103/ijmr.ijmr_1821_22.
[39] R.O. Ogundokun, R. Maskeliunas, S. Misra, R. Damaševičius, Improved CNN based on batch normalization and adam optimizer, in: International Conference on Computational Science and Its Applications Cham, Springer International Publishing.. 2022, pp. 593–604, https://doi.org/10.1007/978-3-031-10548-7_43.
[40] B. Li, F. Wu, S. Liu, J. Tang, G. Li, M. Zhong, X. Guan, CA-Unet++: an improved structure for medical CT scanning based on the Unet++ Architecture, Int. J. Intell. Sys. 37 (2022) 8814–8832, https://doi.org/10.1007/s00521-023-09374-3.
[41] D. Meedeniya, S. De Silva, L. Gamage, U. Isuranga, Skin cancer identification utilizing deep learning: a survey, IET Image Process. 18 (2024) 3731–3749, https://doi.org/10.1049/ipr2.13219.
[42] I. Pacal, B. Ozdemir, J. Zeynalov, H. Gasimov, N. Pacal, A novel CNN-ViT-based deep learning model for early skin cancer diagnosis, Biomed Signal Process. Control 104 (2025) 1–14, https://doi.org/10.1016/j.bspc.2025.107627.
[43] G.H. Dagnaw, M. El Mouhtadi, M. Mustapha, Skin cancer classification using vision transformers and explainable artificial intelligence, J. Med. Artif. Intell. 7 (2024) 1–17, https://doi.org/10.21037/jmai-24-6.
[44] A. Mahbod, R. Ecker, R. Woitek, Fusion of foundation and vision transformer model features for dermatoscopic image classification, arXiv preprint, arXiv:2505.16338 (2025) 7–16, https://doi.org/10.48550/arXiv.2505.16338.
[45] J. Aftab, M.A. Khan, S. Arshad, S.U. Rehman, D.A. AlHammadi, Y. Nam, Artificial intelligence-based classification and prediction of medical imaging using a novel framework of inverted and self-attention deep neural network architecture, Sci. Rep. 15 (2025) 8724–8751, https://doi.org/10.1038/s41598-025-93718-7.
[46] V. Venugopal, N.I. Raj, M.K. Nath, N. Stephen, A deep neural network using modified EfficientNet for skin cancer detection in dermoscopic images, Decis. Anal. J. 8 (2023) 1–12, https://doi.org/10.1016/j.dajour.2023.100278.
[47] M.U. Ali, M. Khalid, H. Alshanbari, A. Zafar, S.W. Lee, Enhancing skin lesion detection: a multistage multiclass convolutional neural network-based framework, Bioengineering 10 (2023) 1430–1456, https://doi.org/10.3390/bioengineering10121430.

[48] V. Ravi, T.J. Alahmadi, T. Stephan, P. Singh, M. Diwakar, DEEPSCAN: integrating vision transformers for advanced skin lesion diagnostics, Open Dermatol. J. 18 (2024) 12–25, https://doi.org/10.2174/1874372224666240212131927.
[49] I. Pacal, B. Ozdemir, A robust deep learning framework for multiclass skin cancer classification, Sci. Rep. 15 (2025) 4938–4957, https://doi.org/10.1038/s41598-025-89230-7.
[50] C. Shorten, T.M. Khoshgoftaar, A survey on image data augmentation for deep learning, J. Big Data 6 (2019) 60–118, https://doi.org/10.1186/s40537-019-0197-0.
[51] K. Zhang, H. Ying, H.N. Dai, L. Li, Y. Peng, K. Guo, H. Yu, Compacting deep neural networks for Internet of Things: methods and applications, IEEE Internet Things J. 8 (2021) 11935–11959, https://doi.org/10.1109/JIOT.2021.3063497.
[52] Z.H. Zhou, Ensemble learning, in: Machine Learning, Singapore Springer Singapore. 2021, pp. 181–210, https://doi.org/10.1007/978-981-15-1967-3_8.
[53] G. Varoquaux, V. Cheplygina, Machine learning for medical imaging: methodological failures and recommendations for the future, NPJ Digit. Med. 5 (2022) 48–56, https://doi.org/10.1038/s41746-022-00592-y.
[54] T.J. Bradshaw, Z. Huemann, J. Hu, A. Rahmim, A guide to cross-validation for artificial intelligence in medical imaging, Radiol. Artif. Intell. 5 (2023) 16–31, https://doi.org/10.1148/ryai.220232.
[55] L. Gamage, U. Isuranga, D. Meedeniya, S. De Silva, P. Yogarajah, Melanoma skin cancer identification with explainability utilizing mask guided technique, Electronics 13 (2024) 680–701, https://doi.org/10.3390/electronics13040680.