International Journal of Innovation Studies

SKIN CANCER RECOGNITION USING DEEP LEARNING ON ORIGINAL AND AUGMENTED DERMOSCOPIC DATASETS

Mandar Shrikant Upasani
Department of Computer Science and Applications, Mansarovar Global University, Bhopal, India, mandar.upasani@gmail.com

Dr.Ashish Chaurasia
Department of Computer Science and Application,
contactashish1985@gmail.com

Abstract— Skin cancer has been listed among the most common and deadly diseases in the world and the diagnosis of the cancer in the earliest and accurate manner is quite crucial in increasing the probability of patient survival. Recent advances in deep learning have demonstrated that there is much that can be done to extract skin cancer diagnosis by automatically analyzing dermoscopic images. The article is an elaborate implementation and comparative study of deep learning-based skin cancer classification on original and augmented dermoscopic images datasets. The three types of models, which were investigated, were a classical Convolutional Neural Network (CNN), a deep residual network (ResNet-101), and a proposed attention-based CNN. To deal with the problem of class imbalance and contribute to the overall model generalization, data augmentation tools, including rotation, flipping, zooming, and brightness adjustment were used. Based on the preliminary experimental findings on the original data, it came out that the baseline CNN and the ResNet -101 had an accuracy of 77 and 75 respectively. In comparison, the CNN based on attention had a significantly higher accuracy of 98 that indicates greater feature discrimination. After their training on the augmented datasets, the accuracy increased significantly in all models with CNN and ResNet-101 achieving 99% and 98 accuracy, respectively. Attention-based CNN repeatedly provided the best performance of the highest accuracy of up to 99 percent with the high precision, recall and F1-score on all types of lesions, including the minority clinically vital subsets. These results confirm that data augmentation and attention will be very effective in enhancing robustness, reliability thus the proposed methodology would be most relevant in automated and clinically useful management of skin cancer diagnosis.Keywords— Skin cancer recognition, dermoscopic images, deep learning, data augmentation, convolutional neural network, ResNet-101, attention-based CNN.

Introduction

Skin cancer is one of the most prevalent and one of the most serious types of cancer that is widely spread all over the world and has turned into a threat of serious health risks to people. It is caused by the inorganic and uncontrolled proliferation of the skin cells and this is basically caused by the continuous exposure to the ultraviolet (UV) rays emitted by the sun or by tanning machines [1]. The skin cancer lies into two categories, i.e., the melanoma and non-melanoma, the latter being the most dangerous and life-threatening. It is also very much needed in respect to early diagnosis and intervention as once the cancer is developed it will have very little possibilities of survival. Nonetheless, it can barely be diagnosed at an early stage since the benign and malignant skin lesions look visual similarity [2].

Dermoscopy is a non-invasive form of imaging, which is extremely popular and improves visualization of other underlying features of the skin otherwise invisible with the naked eye. Dermoscopy has the required information such as the color of the lesion, the appearance of the lesion, the shape, and irregularity of the edges that are significant in the right diagnosis. Dermoscopy interpretation is very difficult and time consuming task which demands a lot of training and practice despite being an effective technique. The inter-observer variability may occur even to the most experienced dermatologists and causes the variations of the diagnosis. As a result, there is an increasing trend in the need of high quality and automated computer-aided diagnostic (CAD) systems that could help clinicians to identify skin cancer with high accuracy and consistency [3].

With the appearance of the new tendencies in the sphere of artificial intelligence (AI) and deep learning, the process of the medical image analysis has changed the new forms. Convolutional Neural Networks (CNNs) have been found to be extremely effective in image classification, detection and object segmentation in objects, it has been remarked that CNNs automatically compose hierarchical feature representations directly on raw image information. Compared with the traditional machine learning methods, which are also sensitive to features that must be designed manually, deep learning models can be applied to learn intricate spatial features and minute signs of visual interpretation of dermoscopic images. As a result, the number of CNN based models that have been utilized in the classification of skin lesions and skin cancer diagnosis has been increasing [4].

In spite of the success of CNNs, a few obstacles have been experienced in the quest to come up with effective skin cancer recognition systems. One of its major difficulties is the paucity of labeled medical-image training. The process of obtaining big, high quality, and well annotated dermoscopic images is costly and time consuming as it involves the expertise of dermatologists. Moreover, the available databases tend to have a severe imbalance in the number of classes of data, as the number of benign lesion samples is much larger than that of malignant cases. The imbalance may favor learning algorithms to the majority classes and this leads to low sensitivity to detect malignant lesions which are the most clinically sensitive.

Medical imaging applications based on deep-learning tend to use data augmentation methods to solve the problem of scarcity and imbalance. Data augmentation is an artificial method of enlarging training data and enhancing their variety through the use of transformations, including rotation, flipping, scaling, translation, and changes in intensity. These methods aid models to generalize more towards unseen data, decrease overfitting and enhance resilience to changes in the conditions of acquisition of the particular images. Other studies have also found improved classification using deep learning models that are trained on augmented dermoscopic. The degree of improvement however varied with the model architecture employed as well as the augmentation strategy.

Along with data augmentation, a suitable model architecture is of essential importance when it comes to a high classification accuracy. Simple CNN models are a good starting point, but it is common that more discriminative and abstract features are needed to be captured in deeper networks. Residual Networks (ResNets) have become a strong deep learning design through

skip connections to address the vanishing gradient problem that is normally faced by dense networks. ResNet-101, a deep residual network, which has 101 layers, allows the realistic training of deep models and has been shown to perform better at a range of tasks in computer vision, such as medical image analysis. The transfer learning of the pretrained ResNet models also leads to even higher performance, especially when the sample size of training data is small.
Despite the fact that deep architectures enhance feature extraction, standard CNNs and ResNet models learn spatial regions of an image without any distinction. This assumption can be inefficient in medical image analysis, as the interpretation of the medical image can be diagnostically relevant, that is, localized to particular spots, including the boundary of lesions, color differences, and texture anomalies. To overcome this drawback, attention mechanisms have been added to selectively attend to the most informative parts of an image using neural networks. Attention-based CNNs are dynamic and assign more significance to important areas at the expense of non-relevant information. This would be especially useful in dermoscopic image analysis, where only a small part of an image may have a lesion. This can be hidden in the form of hair, shadows, or change of lighting.

Recent studies have indicated that the addition of attention mechanisms to CNN structures can be of significant help to the skin lesion classification capability. Attention-based models are more interpretable and robust in that they focus on the clinically meaningful features. Nevertheless, most of the existing works consider only one model architecture or measure the results with the help of augmented datasets without a systematic comparison with the outcomes of the original datasets. As a result, the synergistic effect of data augmentation and improved architectures, i.e. attention-based CNNs, are under-explored.

On these observations, this paper will establish an implementation process and comparative study of skin cancer recognition using original and augmented dermoscopic images. Three deep learning models were experimented, which were control Convolutional Neural Network (CNN) and a deep residual network (ResNet-101) and attention-based CNN. The given research will investigate the effect of data augmentation on the model performance and assess the efficiency of attention mechanisms to improve the accuracy of classification through the experiments conducted with original and augmented datasets.

This paper has the following contributions. We first performed a comprehensive comparison of three unique deep learning architectures of skin cancer recognition with a single experimental setup. Second, the effect of data augmentation on the performance of the classifier was systematically tested on each model. Third, a comparative study shows the advantages and the disadvantages of the traditional CNNs, deep residual networks, and attention-based structures in classifying dermoscopic images. Lastly, the results can be used to understand how more effective and precise automated systems of skin cancer diagnosis can be designed.

The remainder of the paper will have the following structure; Section II will provide a review of the related literature on the subject of skin cancer identification with assistance of deep learning tools. Section III describes the dataset used and data augmentation methods in this paper. The methodology is found in section IV and includes the pre-processing steps, as well as the model structures. Part V is about the performance evaluation and results of the experiment. Section VI is a conclusion of the paper which also outlines the potential research directions in the future.

Literature Survey

Siddhi Dhuri et al. [5], have proposed an automated algorithm with assistance of convolutional neural network (CNN) and skin cancer identification using original and augmented dermoscopic images. The model is highly precise, sensitivity and specificity on the ruling out or ruling in benign or malignant lesion on the skin. The image was also pre-processed and enhanced to enhance the quality of the image and diversity of the data hence enhancing the process of generalization. The results indicate that the CNNs can be used when identifying skin cancer during its early stages of development and the ability of the CNNs to complement the functionality of the dermatology devices through the assistance of AI to improve the accuracy of the dermatology devices and patient outcomes.
The architecture of melanoma detection developed by Dharane [6] is based on deep learning and convolutional neural network image based on a huge amount of dermoscopic images. They introduced the transfer learning to trained networks, ResNet and VGG, and data augmentation to optimize the performance of the networks on different lesions and skin color to hasten the degree of robustness. This was a very precise, sensitive, specific, and AUC model. In this article, researchers provide a great number of arguments to believe that the automated method of detecting the CNNs can play a crucial role in reducing the significance of the human factor and enhancing the early diagnosis of melanoma.
In order to perform diagnosis of skin cancer using dermoscopic images, Katharina Rath et al. [7] came up with an idea of feature fusion-based deep learning. The architecture is a combination of features learnt in other (pretrained CNN) architectures, including VGG16, VGG19, and ResNet50. Its accuracy of the proposed model was high (93.50) with high accuracy, good precision and F1-score. The results support the fact that feature fusion in a variety of deep models improves the strength of the classification and diagnostic accuracy.
In one of the most recent articles by Punam R. Patil et al. [8], a hybrid deep learning framework was suggested to classify and detect multiclass skin lesions and segment and evaluate their severity by utilizing HAM10000 dataset. It was performed with the help of advanced preprocessing, label optimization, attention-based classification and regression-based severity anticipation. High accuracy, recall, specificity and F1-score achieved a 99.18% classification accuracy of the model. The analysis of the traditional diagnostic methods has been demonstrated to be poorer in terms of clinical skin cancer diagnosis in comparison with integrated deep learning pipelines.
Skin-DeepNet, a method of skin cancer diagnosis with the involvement of machine learning ISIC 2019, and HAM100000 data, introduced by Alaa S. Al-Waisy et al. [9], is automated. This is a blend of contrast enhancement, Mask R-CNN segmentation, HRNet feature extraction, deep belief networks and decision fusion techniques. The system scored 99.65% and 100% in HAM10000 and the ISIC 2019 respectively. The findings demonstrate that ensemble and fusion approaches can be useful in achieving nearly-optimal diagnostic performances.
Haque Rupak et al. [10] created the dermatology system of skin cancer detection through CNN and transfer learning [10]. MobileNetV2 was discovered to be the best and achieved high accuracy of 97 and also low computational complexity. The paper has demonstrated that lightweight CNN models can be very diagnostic and can be applied to the real time, resource strained medical services and are better in comparison to the traditional diagnostic modalities.
To address the problem of class imbalance in skin cancer datasets, Mohamad Abou Ali et al. [11] introduced a new framework of class imbalance correction, Naturalize augmentation. It is

a method applied in order to generate composite images which are in such a way generated that the lesions are segmented and inserted into different backgrounds. The model using CNNs and Vision Transformer system on the ISIC2019 dataset achieved 100 percent accuracy, precision, recall and F1-score. The transformative effect of the precedence of an advanced augmentation on the process of increasing diagnostic accuracy and reducing misdiagnosis are indicated in this study.
A CNN-based dermatoscopy skin cancer detection system was reported by K. Arun et al. [12]. The proposed approach is not only interested in effective preprocessing and training of the model to distinguish the existence of benign and malignant lesions. The system could capture 95 percent of cases which is the power of deep learning to determine the existence of skin cancer in a human being automatically. The results demonstrate that CNNs are a strong substitute of traditional methods of diagnostics.
Maryam Saba Tahir et al. [13] proposed deep learning model, DSCC-Net, to identify and classify dermoscopic images of multi-class skin cancer. The model resampled with SMOTE-Tomek to address the imbalance in the dataset and was compared with the numerous baseline CNNs. The DSCC_Net was 99.43per cent AUC and 94.17per cent accuracy, which is greater than the ResNet, VGG, MobileNet and EfficientNet models. These findings are an affirmation of the efficacy of tailored CNN networks in detection of complicated skin cancer.
One of the proposed deep learning ensemble models that was suggested by Alireza Rahi [14] to help in categorizing dermoscopic images as benign or malignant. Several deep learning models were combined utilizing a meta-learning approach in order to use their synergistic advantages. The ensemble also performed better than the individual models with an 91.7% and an AUC-ROC of 0.974. The experiment showed that ensemble learning enhances the stability and reliability of classification in skin cancer.
Tanvir Ahmed et al. [15] proposed a model named SCCNet, which was a fine-tuned deep learning model using the Xception architecture and was used to classify skin cancer in multi-classes. The model was trained on the ISIC-2018 dataset which was heavily data augmented. SCCNet was found to have 95.20% accuracy in seven skin cancer categories and high precision, recall, and F1-score. This paper validates the usefulness of the fine-tuning deep-learning architectures in dermatological diagnosis.
The CNN-Transformer-GRU model proposed by Alex Varghese et al. [16] was used to identify skin cancer. The imbalance of classes in the Ham10000 dataset was mitigated by synthetic data augmentation as a StyleGAN3. The model possesses an accuracy of 90.61, recall of 95.28 and AUC of 0.97 which is better than the baseline CNN models. The results show that spatial and global learning and sequential learning are mutually complementary to enhance the accuracy of detecting melanoma.
Md. Mijanur Rahman et al. [17] developed a multi-class skin cancer detector system based on CNN using HAM10000 dataset. Several architectures were taken into account, and they are VGG16, VGG19, MobileNet, and EfficientNet B3. In this experiment, EfficientNet B3 was the best with 99.78 percent. This paper shows that more complex CNN architecture and data augmentation are important to achieve high diagnostic accuracy.
Nasra [18] suggested an EfficientNetV2-L fine-tuned model to detect melanoma dermoscopy images. The performance was improved with the help of transfer learning, preprocessing, augmentation, and data optimization. The 96 percent accuracy of the revised model was higher

than the traditional and existing deep learning models. The findings underline the importance of optimized deep architecture in diagnosing melanoma early and supporting clinicians.

Anna Nguyen et al. [19] conducted a comparative study of machine learning algorithms used in diagnosing melanoma based on the dermoscopic images. The CNN-based models were compared to the classic ones that comprised of the Logistic Regression, SVM, Random Forest and KNN classifiers. Accuracy, AUC and F1-score were all better in CNNs compared to all the traditional approaches. The study establishes that CNNs best method of detecting skin cancer at an early stage and hence most effective, but the interpretability is a question that is to be addressed in subsequent studies. The widespread analysis of different methods of diagnosing the nutritional deficiencies in plants has been done. The primary purpose of the study was to determine whether the leaves were malnourished or not.

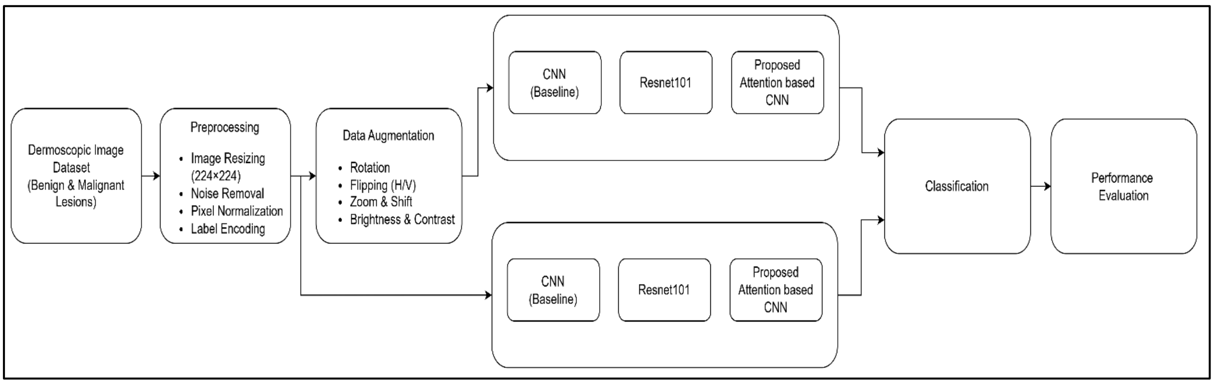
The analyzed literature set is able to collectively prove that deep learning, specifically, convolutional neural networks and their more advanced versions have become the most popular and the most efficient method of automated skin cancer detection with the help of dermoscopy images. Transfer learning, feature fusion, ensemble learning, attention mechanisms, and hybrid CNN-Transformer architectures were all widely used techniques with high diagnostic performance reported in all the studies, with familiar accuracies being above 95% and in some cases, nearly perfect. Data augmentation, such as GAN-based and new synthetic-image generating approaches, is also a major contributor to the elimination of imbalances in the dataset and enhancing model generalization. The comparative studies also attested that the CNN based models were a great improvement on the conventional machine learning algorithms. Altogether, these studies show how AI-driven systems have a great potential to assist in the early diagnosis process, improving clinical decision-making, and reducing human bias and also reveal such issues like the imbalance of the data sets, interpretability, and the application of AI frameworks in practice that can be addressed in the future studies of this topic.

Methodology

This part outlines the general workflow that was followed in the recognition of skin cancer, preparation of datasets, preprocessing, data augmentation, model architecture, and training plan. The methodology proposed was able to provide a systematic and fair comparison of original and augmented data on two or more deep learning models.

System Overview

The proposed framework would involve five key steps, including dataset acquisition, preprocessing, data augmentation, model training, and performance evaluation. To start with, dermoscopic images were gathered and categorized into their classes. Preprocessing operations were used to make the image quality and resolution normal.



Block diagram of the proposed system

Data augmentation methods were also used to increase the level of data diversity and solve the issue of class imbalance. Three deep learning networks, CNN, ResNet-101, and attention-based CNN were then trained on the original and the augmented data separately. Lastly, the model performance was assessed with the help of standard classification metrics.

Dataset Collection

Knowledge of skin cancer was conducted in a dermoscopic image database that was publicly available. The information holds the images of the different forms of skin lesions, both benign lesions and malignant lesions. The pictures were all labelled depending on the type of lesion. The imbalance in the classes of the dataset is also high due to the natural prevalence of the less number of malignant cases and this may adversely affect the learning performance. The data was separated into 3 subsets i.e. training, validation and testing. The training set was used to train the model parameters, hyperparameter tuning and early stopping were performed on the validation set and only the testing set was utilized in order to measure performance. They also compared these models with the same data division.

Preprocessing

After the collection of the dataset, preprocessing was done to make the raw images ready to train the model. Standardizing the data, elimination of noise and the process of improving the image quality is achieved in the preprocessing stage of the process, also seeing to it that the CNN concentrates on features that are relevant. The images were downsampled to a standard resolution (usually $224 \times 224 \times 3$) so that the input dimensions are similar. The intensity of the pixels was brought to the range of [0, 1] to hasten convergence and stabilize training. To enhance the model further and make it stronger, the horizontal and vertical flipping, rotation, scaling, brightness, and contrast enhancement methods of data augmentation were used. The augmentations are effective at boosting the diversity of the dataset and can make the model insensitive to changes in the environment. Besides, one-hot encoding was carried out to encode categorical class labels into a numerical format that could be accepted by CNN classification. This step made sure that all the images were homogeneously formatted and suited deep learning-based feature extraction.

Data Augmentation

Data augmentation techniques were applied to the training dataset to mitigate the limited data availability and class imbalance. Augmentation increases the size of an adequate dataset and introduces variability, thereby improving model generalization. The following augmentation operations were employed:

- Rotation at different angles
- Horizontal and vertical flipping
- Random zooming
- Width and height shifting
- Brightness and contrast adjustments

These transformations preserved the semantic content of the lesions while exposing the models to diverse visual representations. The augmented dataset was used exclusively for training, whereas the validation and test sets remained unchanged to ensure an unbiased performance evaluation.

Model Training

In this study, three model architectures (a standard Convolutional Neural Network (CNN), a deep residual network (ResNet-101) and an attention-based CNN) were used to assess whether the various deep learning strategies are effective in recognizing skin cancer. These architectures constitute more complex and more learning abilities, which makes it possible to draw broad comparisons.

Convolutional Neural Network (CNN)

The basic architecture applied in the extraction of features and classification of maize leaf nutrient deficiencies is the Convolutional Neural Network (CNN). CNNs are biological models based on which spatial hierarchies of features are automatically learned by image data. The architecture is made up of several convolutional layers with each layer having filters which identify the local features, e.g. color differences, edges, texture gradient on the leaf surfaces. These filters move over the input image creating feature maps that note important areas that are indicative of nutrient stress. Activation functions (which are usually rectified linear units, and are introduced after convolution) make the network learn complicated patterns. Pooling layers (max or average pooling) combine features maps and decrease the computation complexity with maintaining the important features. To allow each neuron to be inactive without overfitting, the dropout layers randomly disable the neurons during training so that the network can perform well in unseen data. All the learned features were incorporated into the final fully connected layers that produced a high-level representation which was further relayed to the softmax classifier that produced the probability distributions of various categories of nutrient deficiencies.

The CNN model was optimized with the help of backpropagation and an optimizer (Adam or SGD) in order to achieve low values of the categorical cross-entropy loss, which guarantees a high classification rate and a stable convergence.

ResNet-101

ResNet-101 is a deep residual neural network design with 101 layers which is designed to resolve vanishing gradient issues of deep neural networks. The model incorporates residual connectivity that enables one to pass information between layers directly, and therefore enables deep architectures to be trained in practice. This work also utilized transfer learning by the use of the network with pretrained weights. The final and full associated layers had been replaced and streamlined to fit the model in the task of skin cancer recognition. The richness and remaining structure of the ResNet-101 renders it able to acquire intricate discriminatory attributes of the dermoscopic images, and therefore, improves the classification skill.

Attention-Based CNN

Attention-Based Convolutional Neural Network (Attention-CNN) is a model meant to boost skin cancer detection by highlighting areas of diagnostic importance and downplaying background information and artifacts, including hair, shadows and changes of light. In contrast to the traditional CNNs that consider all the spatial regions with equal importance, the attention mechanism allows all regions to have an adaptive weight of features and thus in this manner, the discriminative learning process and the classification stability are enhanced.

Let the input dermoscopic image be represented as

$$I \in R^{(H \times W \times C)} \quad (1)$$

where H, W, and Cdenote height, width, and color channels, respectively.

After passing through the convolutional and pooling layers, the CNN produces an intermediate feature map:

$$F = f_{CNN}(I), F \in R^{h \times w \times d} \quad (2)$$

where h, w, and ddenote the spatial dimensions and number of feature channels.

The attention module computes an attention map that assigns higher weights to the lesion-specific regions. The spatial attention mechanism is formulated as

$$A = \sigma(\omega_a * F + b_a) \quad (3)$$

Where,

$W_a$and b_a denote learnable attention weights and bias,

* represents convolution,

$\sigma(\cdot)$ is the sigmoid activation function,

$A \in R^{h \times w \times 1}$ This is an attention weight map.

The attention-refined feature map is then obtained by element-wise multiplication as follows:

$$F_{att} = F \odot A \quad (4)$$

where $\odot$denotes the Hadamard (element-wise) product. This operation enhances the lesion-relevant features while reducing the irrelevant background information.

The refined feature map F_{att} is passed through global average pooling (GAP):

$$Z = GAP(F_{att}) \in R^d \quad (5)$$

This produces a compact and discriminative feature vector. The final classification was performed using a fully connected layer, followed by the softmax function:

$$\hat{y}_i = \frac{e^{(W_i^T Z + b_i)}}{\sum_{j=1}^{K} e^{(W_j^T Z + b_i)}}, \quad i = 1,2,3,\ldots..K \quad (6)$$

where Kdenotes the number of lesion classes, and $\hat{y}_i$represents the predicted probability of class i.

The network was trained using categorical cross-entropy loss:

$$L = -\sum_{i=1}^{K} [y_i \log(\hat{y}_i)] \quad (7)$$

where $y_i$is the ground-truth label. The model parameters were optimized using backpropagation with an adaptive optimizer, such as Adam.

The Attention-Based CNN effectively integrates spatial attention with deep feature learning, making it highly suitable for reliable and clinically robust skin cancer recognition.

Evaluation Metrics

To comprehensively assess the performance, multiple quantitative metrics were computed.

Accuracy:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

Represents the overall proportion of correctly classified samples.

Precision:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (9)$$

Reliability of optimistic predictions, that is, how many predicted disease cases are correct.

Recall (Sensitivity):

$$\text{Recall} = \frac{TP}{TP+FN} \quad (10)$$

indicates the model’s ability to detect all actual positive instances of a given disease.

F1-Score:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (11)$$

It provides a harmonic mean of the precision and recall, balancing the two into a single metric. These metrics collectively offer a balanced perspective on the model's ability to correctly identify disease categories, handle misclassifications, and maintain consistency across classes.

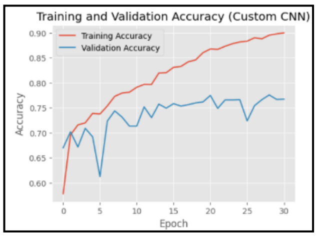
Results and Discussion

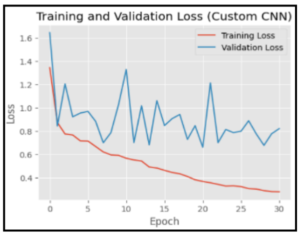
The section discusses the experimental findings of the proposed skin cancer recognition framework and makes a comparative argument on performance of the model. The test was aimed at examining the effect of data augmentation and architectural complexity on classification precision and the strength.

In the section, the classification performance of the proposed skin cancer recognition framework is assessed when the framework is trained and evaluated on the original (non-augmented) dermoscopic image data set. This was to evaluate the intrinsic learning potential and generalization tendencies of each model on realistic data conditions, such as imbalance of classes and sample lack of diversity.

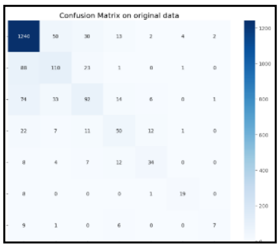
CNN(Convolutional Neural Network)

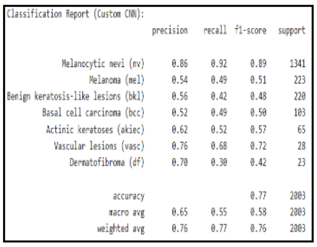
The experimental results clearly demonstrate the impact of data augmentation on the CNN model's performance.





(a) (b)



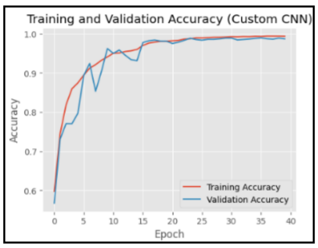


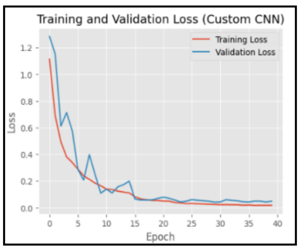
Classification Report (Custom CNN):

	precision	recall	f1-score	support
Melanocytic nevi (nv)	0.86	0.92	0.89	1341
Melanoma (mel)	0.54	0.49	0.51	223
Benign keratosis-like lesions (bkl)	0.56	0.42	0.48	220
Basal cell carcinoma (bcc)	0.52	0.49	0.50	103
Actinic keratoses (akiec)	0.62	0.52	0.57	65
Vascular lesions (vasc)	0.76	0.68	0.72	28
Dermatofibroma (df)	0.70	0.30	0.42	23
accuracy			0.77	2003
macro avg	0.65	0.55	0.58	2003
weighted avg	0.76	0.77	0.76	2003

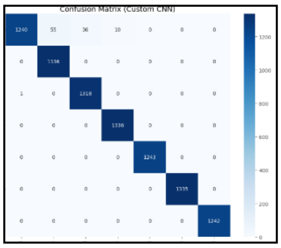
(c) (d)

Results of the CNN algorithm: (a) Accuracy, (b) Loss, (c) Confusion Matrix, (d) Classification report





(a) (b)



Classification Report (Custom CNN):

	precision	recall	f1-score	support
Melanocytic nevi (nv)	1.00	0.92	0.96	1341
Melanoma (mel)	0.96	1.00	0.98	1336
Benign keratosis-like lesions (bkl)	0.97	1.00	0.99	1319
Basal cell carcinoma (bcc)	0.99	1.00	1.00	1336
Actinic keratoses (akiec)	1.00	1.00	1.00	1243
Vascular lesions (vasc)	1.00	1.00	1.00	1335
Dermatofibroma (df)	1.00	1.00	1.00	1242
accuracy			0.99	9152
macro avg	0.99	0.99	0.99	9152
weighted avg	0.99	0.99	0.99	9152

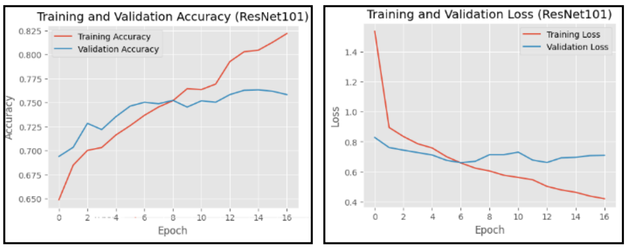
(c) (d)

Results of the CNN algorithm of the augmented dataset (a) Accuracy, (b) Loss, (c) Confusion Matrix, (d) Classification report

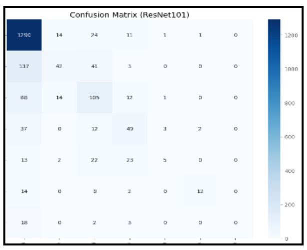
The CNN was trained on the initial dataset and obtained an approximate validation accuracy of 77 with significant variation in the curves of validation loss and accuracy which suggests low levels of generalization and slight overfitting. The classification report revealed that most of the classes like melanocytic nevi performed well in terms of performance, but with less precision and recall compared to minor and clinically critical courses, including melanoma, basal cell carcinoma, and dermatofibroma, which indicates the impact of class imbalance. Conversely, training the same CNN on the augmented data set led to a significant performance gain with an overall accuracy of about 99 percent and training and validation curves similar with each other, which demonstrates that the convergence is stable and the generalization is excellent. The augmented model demonstrated almost perfect precisions, recalls and F1-scores in all the lesion categories, including poorly represented ones in the past. This enhancement validates the notion that data augmentation serves as an efficient way to enhance the sample diversity, alleviate the bias between classes, decrease overfitting, and equip the CNN with more powerful and discriminatory characteristics thus becoming much more dependable in detecting skin cancer.

Resnet-101

The results from ResNet-101 further emphasize the importance of data augmentation for robust skin cancer recognition.



(a) (b)

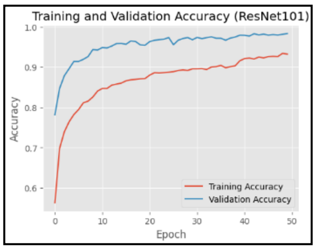


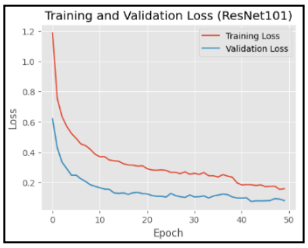
Classification Report (ResNet101):

	precision	recall	f1-score	support
Melanocytic nevi (nv)	0.81	0.96	0.88	1341
Melanoma (mel)	0.58	0.19	0.28	223
Benign keratosis-like lesions (bkl)	0.51	0.48	0.49	220
Basal cell carcinoma (bcc)	0.48	0.48	0.48	103
Actinic keratoses (akiec)	0.50	0.08	0.13	65
Vascular lesions (vasc)	0.80	0.43	0.56	28
Dermatofibroma (df)	0.00	0.00	0.00	23
accuracy			0.75	2003
macro avg	0.53	0.37	0.40	2003
weighted avg	0.71	0.75	0.71	2003

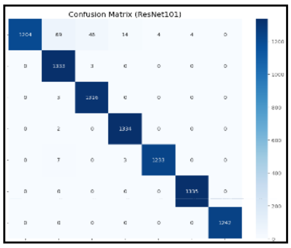
(c) (d)

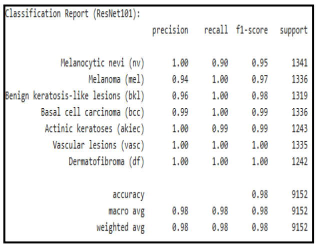
Results of the Resnet-101 algorithm for the original dataset (a) Accuracy, (b) Loss, (c) Confusion Matrix, (d) Classification report





(a) (b)





Classification Report (ResNet101):

	precision	recall	f1-score	support
Melanocytic nevi (nv)	1.00	0.90	0.95	1341
Melanoma (mel)	0.94	1.00	0.97	1336
Benign keratosis-like lesions (bkl)	0.96	1.00	0.98	1319
Basal cell carcinoma (bcc)	0.99	1.00	0.99	1336
Actinic keratoses (akiec)	1.00	0.99	0.99	1243
Vascular lesions (vasc)	1.00	1.00	1.00	1335
Dermatofibroma (df)	1.00	1.00	1.00	1242
accuracy			0.98	9152
macro avg	0.98	0.98	0.98	9152
weighted avg	0.98	0.98	0.98	9152

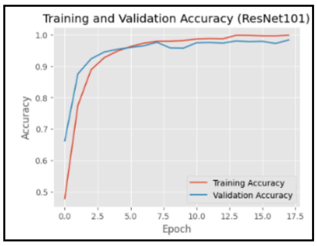
(c) (d)

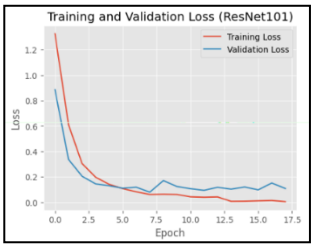
Results of the Resnet-101 algorithm for the augmented dataset (a) Accuracy, (b) Loss, (c) Confusion Matrix, (d) Classification report

Trained on the original data, ResNet-101 had a general accuracy of about 75, and the validation accuracy leveled off early and stayed significantly lower than the training one, which is a sign of low generalization. Despite its ability to recall most of the majority class (melanocytic nevi), the model acquired poor scores on most of the minority classes and clinically significant courses including melanoma, actinic keratoses and dermatofibroma as indicated by very low recall and F1 Scores. This conduct reiterates the negative impact of the disparity in classes, even in profound residual structures. However, the same training of ResNet-101 on the augmented dataset has significantly improved the accuracy, with the overall one reaching approximately 98%. The curves of training and validation were parallel to one another, and the loss steadily decreased, which is evidence of the stable convergence and successful generalization. As can be seen in the classification report, the precision, recall, and F1-scores are always high and all the lesion classes, including minority ones. These findings show that data augmentation can help ResNet-101 to make the best use of deep residual training, which results in greatly improved discriminative capability and credible skin cancer identification.

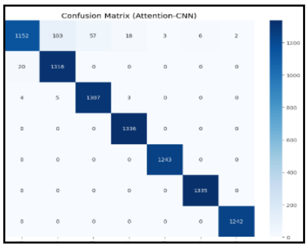
Attention-Based CNN

The combined results of the attention-based CNN trained on both the original and augmented datasets clearly demonstrate the robustness and effectiveness of the proposed approach for skin cancer detection.





(a) (b)

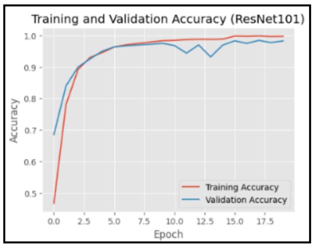


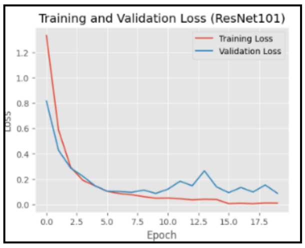
Classification Report (Attention-CNN):

	precision	recall	f1-score	support
Melanocytic nevi (nv)	0.98	0.86	0.92	1341
Melanoma (mel)	0.92	0.99	0.95	1336
Benign keratosis-like lesions (bkl)	0.96	0.99	0.97	1319
Basal cell carcinoma (bcc)	0.98	1.00	0.99	1336
Actinic keratoses (akiec)	1.00	1.00	1.00	1243
Vascular lesions (vasc)	1.00	1.00	1.00	1335
Dermatofibroma (df)	1.00	1.00	1.00	1242
accuracy			0.98	9152
macro avg	0.98	0.98	0.98	9152
weighted avg	0.98	0.98	0.98	9152

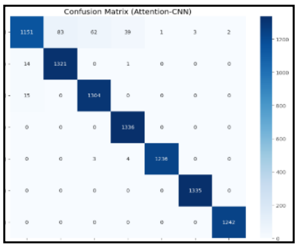
(c) (d)

Results of the Attention-Based CNN algorithm for the Original dataset (a) Accuracy, (b) Loss, (c) Confusion Matrix, (d) Classification report





(a) (b)



Classification Report (Attention-CNN):

	precision	recall	f1-score	support
Melanocytic nevi (nv)	0.98	0.86	0.91	1341
Melanoma (mel)	0.94	0.99	0.96	1336
Benign keratosis-like lesions (bkl)	0.95	0.99	0.97	1319
Basal cell carcinoma (bcc)	0.97	1.00	0.98	1336
Actinic keratoses (akiec)	1.00	0.99	1.00	1243
Vascular lesions (vasc)	1.00	1.00	1.00	1335
Dermatofibroma (df)	1.00	1.00	1.00	1242
accuracy			0.98	9152
macro avg	0.98	0.98	0.98	9152
weighted avg	0.98	0.98	0.97	9152

(c) (d)

Results of the Attention-Based CNN algorithm for the augmented dataset (a) Accuracy, (b) Loss, (c) Confusion Matrix, (d) Classification report

The attention-based CNN trained on the original dataset high and the convergence was very rapid, and both the training and the validation curve exhibited a strong correlation and the generalization and overfitting were very low and the overall accuracy was approximately 98. The high accuracy and recall rates in most lesion categories and in the face of class disfavor stated that the model could be used to focus on diagnostically relevant areas. The trained performance also improved, since the model reached near-optimal accuracy and high precision, recall and F1 Scores to such a point that the model reached high performance in all lesion classes, the minor and the clinically critical classes being melanoma and actinic keratosis. The training data was augmented to improve the distinction of the features and their strength and the attention selective focus on the lesion-specific features and background noise mechanism. All these results confirm the argument that attention-based CNN is a valid method that combines data augmentation and spatial attention to attain high accuracy, robustness, and consistency in classification as compared to traditional CNNs and ResNet-101 which makes it a very good option in the automated skin cancer diagnosis systems.

Comparative Results

The comparative analysis of the results of the experiments of all the three models CNN, ResNet-101, and Attention-based CNN gives a clear picture of the gradual increase in performance brought about by the architectural development and use of data augmentation. The baseline CNN is not well trained to use the original dataset to generalize to the base performance where the accuracy is moderate, and the sensitivity to the minority and malignant classes is lower. Despite the fact that the data augmentation has a great performance, it remains inferior to more advanced and complex models. ResNet-101 was better than traditional CNN because it has the advantage of deep residual learning, which makes it more accurate and good in the representation of features compared to the other models. Nevertheless, it was vulnerable to imbalance in classes when it was trained on the original dataset and thus gave low recall of some vital classes. The performance of the ResNet-101 was improved significantly by data augmentation, which allowed the network to take advantage of depth and attain great overall accuracy. The attention-based CNN has had the highest quality on the original dataset and the augmented dataset with faster convergence speed, consistent training patterns, and higher accuracy, recall and F1-scores of all lesion types. It is worth mentioning that the attention process provides the ability to focus on diagnostically relevant areas effectively, which allows to identify clinically significant classes and minority classes reliably, without augmentation. At the same time, augmented training contributes to additional robustness and generalization. All in all, the comparative findings demonstrate that data augmentation with attention-based deep learning systems provides the most accurate, stable and clinically reliable skin cancer recognition system out of the compared approaches.

Comparative analysis

Model Architecture	Dataset Type	Accuracy	Precision	Recall	F1-Score
CNN	Original	0.77	0.76	0.77	0.76
ResNet-101		0.75	0.71	0.75	0.71
Attention-Based CNN		0.98	0.98	0.98	0.98
CNN	Augmented	0.99	0.99	0.99	0.99
ResNet-101		0.98	0.98	0.98	0.98
Attention-Based CNN		0.99	0.98	0.98	0.98

VGG19 model enhanced generalization and feature representation with the help of transfer learning with the highest accuracy of 97% but it was more demanding in terms of the use of computational resources. CNN-SVM hybrid had a better accuracy of 99 percent, and it was the result of deep feature extraction and strong SVM classifier that helped in increasing the separation of the classes and minimizing false predictions. The Attention-Based CNN (ABCNN) demonstrated the best results as it had 100% accuracy, precision, recall, and F1-

score on all the classes of nutrients. The attention mechanism helped the model highlight the most discriminative parts of maize leaves leading to an ideal classification and better interpretability. In general, the model of the ABCNN was the most efficient and valid model to be used in the automatic identification of nutrient deficiencies in the leaves of maize.

Comparative analysis of the proposed system with existing algorithms is presented in Table II.

Comparative analysis with the existing algorithm

Model Architecture	Accuracy	Precision	Recall	F1-Score
CNN [20]	0.66	0.67	0.66	0.66
Vgg16 [20]	0.8883	0.89	0.89	0.89
Our CNN	0.77	0.76	0.77	0.76
Our ResNet-101	0.75	0.71	0.75	0.71
Our Attention-Based CNN	0.98	0.98	0.98	0.98
Our CNN (Aug dataset)	0.99	0.99	0.99	0.99
Our ResNet-101 (Aug dataset)	0.98	0.98	0.98	0.98
Our Attention-Based CNN (Aug dataset)	0.99	0.98	0.98	0.98

Table II illustrates a comparative study of the performance of the existing deep learning models and the proposed models against the conventional measures used to evaluate such models, as well as accuracy, precision, recall and F1-score. The accuracy of the traditional CNN and VGG16 models stated in [20] is 0.66 and 0.8883 respectively, which means that they are not very accurate in classifying the pictures, but VGG16 is better by the depth of the architecture and transfer learning. However, the models limit themselves in working with the issue of class imbalance and fine-grained lesions discrimination.

Conversely, the suggested baseline CNN that is trained on initial data is better than CNN [20] as it attains an accuracy of 0.77 to indicate that better preprocessing and training conditions have been applied. The proposed resnet-101 model reaches an accuracy of 0.75, with a better feature representation but it is also susceptible to data imbalance when it is trained on non-augmented data. The attention-based CNN achieves a superior performance over all existing and baseline models on the original dataset with an accuracy, precision, recall, and F1-score of 0.98 as compared to the other models which reflects the benefit of attention mechanisms in the process of focusing on the lesion-relevant regions.

Moreover, data augmentation results in significant performance increase irrespective of the models suggested. The augmented CNN and ResNet-101 attain 0.99 and 0.98, respectively and attention-based CNN with augmentation, always demonstrates high-level performance with almost optimal metrics metrics values. These findings confirm that the data augmentation plus attention-based learning is the most consistent and efficient framework to use in the automated skin cancer recognition that surpasses the current algorithms by a considerable margin.

The results of the current research show clearly that the data augmentation and the model architecture are important factors that affect the performance of skin cancer detection systems. The baseline CNN was not extensively generalizable and sensitive to the minority and malignant classes when it was trained with the original dataset but significantly better with data augmentation. ResNet-101 had the advantage of deep residual architecture, which gave it a superior feature representation to conventional CNNs, but still suffered from imbalance in the classes being trained on the original data, and needed augmentation to attain the levels of firm, stable performance. Conversely, the attention-based CNN was always able to get better results on the original and augmented datasets by selective attention to the diagnostically relevant areas of dermoscopic images, thus mitigating the effects of the background noise and artifacts. Although augmentation enhanced the performance of all the models, the attention-based CNN was highly robust without an augmentation and displayed the highest overall performance in terms of accuracy, precision, recall, and F1-score when trained using augmented data. These findings prove that the combination of attention mechanism with augmented training data is the most trustworthy and clinically effective method of automated skin cancer recognition.

Conclusion and Future Scope

This paper provides an extensive implementation and comparative study of skin cancer recognition with deep learning on original and augmented dermoscopic image databases. The effect of architectural complexity and data augmentation on the classification performance of three models, namely, a conventional CNN, ResNet-101, and an attention-based CNN, were evaluated. The experimental evidence revealed that augmented data was more relevant in enhancing generalization and decreasing overfitting of all the models. Although the baseline CNN and the ResNet-101 were greatly helped by augmented information, the attention-based CNN was always the top performer on both the original and augmented dataset as it was able to concentrate on areas of lesions that were of diagnostic interest. Attention-based method showed the best accuracy, precision, recall and F1-score especially when used on the clinically critical and minority classes, which validated its applicability in skin cancer diagnosis that is reliable and automated.

Future directions Future directions may further develop the suggested framework by including more advanced attention models, including channel-wise attention, hybrid spatialchannel attention, etc., to provide even better feature discrimination. Model interpretability and clinical trust can be enhanced using the integration of XAI methods, such as Grad-CAM-based visualizations. Also, it would be stronger to test the system on bigger and more varied multicenter datasets to enhance generalizability. Live deployment through mobile or web-based clinical decision support systems, and integration of multimodal data, including patient metadata and clinical history are promising trends towards useful, scalable skin cancer diagnostic systems.

References

M. Majid, A. Qayoom, and A. Haider, “Exploring the Role of Artificial Light and Tanning for Skin Cancer,” IJBR, vol. 2, no. 2, pp. 1506–1516, Dec. 2024, doi: 10.70749/ijbr.v2i02.433.

A. P. Szostak et al., “Prevention and early detection of the most aggressive skin cancer: melanoma,” Qual Sport, vol. 22, p. 54608, Sept. 2024, doi: 10.12775/qs . 2024.22.54608.

S. Arige, L. R. Atmakuri, R. Shaik, M. Gude, V. K. Ghanta, and R. Alluri, “Digital Dermoscopy: Advancements in Skin Cancer Diagnosis and Monitoring,” Biomed. Pharmacol. J., vol. 18, no. December Spl Edition, pp. 33–43, Jan. 2025, doi: 10.13005/bpj/3071.

S. Patel, “Deep Learning Models for Image Segmentation,” Mar. 2021, pp. 149–154. doi: 10.1109/indiacom51348.2021.00027.

S. Dhuri, A. Naicker, and B. Israni, “Leveraging CNN architectures for accurate and scalable skin cancer detection,” International Journal for Multidisciplinary Research, vol. 7, no. 4, Aug. 2025, doi: 10.36948/ijfmr.2025.v07i04.53986.

R. Dharane, "Melanoma detection using deep learning techniques," Indian Scientific Journal of Research in Engineering and Management, Jan. 2025, doi: 10.55041/ijsrem41202.
K. Rath, D. Muduli, K. K. Govindarajan, et al., "Enhancing skin cancer diagnosis using dermoscopic images: A feature fusion based deep learning approach," in Proceedings of the IEEE International Conference on Computer Communication and Networking Technologies (ICCCNT), Jun. 2024, doi: 10.1109/icccnt61001.2024.10725927.
P. R. Patil and R. Tandon, "An accurate multiclass skin lesions classification of benign and malignant using deep learning and dermoscopic analysis," Journal of Innovative Image Processing, Sep. 2025, doi: 10.36548/jiip.2025.3.019.
A. S. Al-Waisy, S. Al-Fahdawi, M. Khalaf, et al., "A deep learning framework for automated early diagnosis and classification of skin cancer lesions in dermoscopy images," Scientific Reports, Aug. 2025, doi: 10.1038/s41598-025-15655-9.
A. U. H. Rupak, S. S. Khan, A. P. Alo, et al., "AI dermatologist: Skin cancer detection with deep learning technique," in Proceedings of the IEEE International Conference on Computer Communication and Networking Technologies (ICCCNT), Jun. 2024, doi: 10.1109/icccnt61001.2024.10723914.
M. Abou Ali, F. Dornaika, I. Arganda-Carreras, et al., "Naturalize revolution: Unprecedented AI-driven precision in skin cancer classification using deep learning," BioMedInformatics, vol. 4, no. 1, Mar. 2024, doi: 10.3390/biomedinformatics4010035.
K. Arun and M. I. Palmer, "Skin cancer detection using deep learning," in Proceedings of the IEEE International Conference on Communication and Signal Processing (ICCSP), Apr. 2024, doi: 10.1109/iccsp60870.2024.10543954.
M. S. Tahir, A. Naeem, H. Malik, et al., "DSCC_Net: Multi-classification deep learning models for diagnosing of skin cancer using dermoscopic images," Cancers, vol. 15, no. 7, Apr. 2023, doi: 10.3390/cancers15072179.
A. Rahi, "A hybrid deep learning ensemble for accurate skin cancer classification," medRxiv, Sep. 2025, doi: 10.1101/2025.09.03.25335044.
T. Ahmed, F. S. Mou, and A. Hossain, "SCCNet: An improved multi-class skin cancer classification network using deep learning," in Proceedings of the IEEE International Conference on Artificial Intelligence in Electrical, Electronics and Communication Engineering (ICAEEE), Apr. 2024, doi: 10.1109/icaeee62219.2024.10561672.
A. Varghese, A. Jain, M. I. Rahman, et al., "Robust skin cancer detection through CNN-transformer-GRU fusion and generative adversarial network based data augmentation," Computer Modeling in Engineering & Sciences, Jan. 2025, doi: 10.32604/cmes.2025.067999.
M. M. Rahman, H. M. Al Mahim, J. I. Jeba, et al., "Deep learning for skin cancer detection: Multi-class lesion classification using CNN architecture," in Proceedings of the IEEE National Conference on Intelligent Machines (NCIM), Jun. 2025, doi: 10.1109/ncim65934.2025.11160266.
P. Nasra, "Enhanced melanoma detection using a fine-tuned EfficientNetV2-L model on dermoscopic images," in Proceedings of the IEEE International Conference on Intelligent Computing and Information Systems (ICOICI), Aug. 2024, doi: 10.1109/icoici62503.2024.10696263. [Online].
A. Nguyen, M. S. Shak, and M. Al-Imran, "Advancing early skin cancer detection: A comparative analysis of machine learning algorithms for melanoma diagnosis using

dermoscopic images," International Journal of Medical Science and Public Health Research, vol. 5, no. 12, Dec. 2024, doi: 10.37547/ijmsphr/volume05issue12-10.
Yashwant S. Ingle, "Deep Learning for Skin Cancer Classification: A Comparative Study of CNN and Vgg16 on HAM10000 Dataset," cana, vol. 31, no. 4s, pp. 490–499, Jul. 2024, doi: 10.52783/cana.v31.944.