Enhanced Melanoma Detection: Leveraging CNN with Convolutional Block and Residual Attention Modules

Gatram Rama Mohan Babu[1], Sivaneasan Bala Krishnan[2], Mohammed Ali Hussain[3], Prasun Chakrabarti[4] [1]Postdoc Scholar,
Singapore Institute of Technology, Singapore.
Professor, RVR & JC College of Engineering, Guntur, India.

[2]Singapore Institute of Technology, Singapore

[3]Professor of CSE, Sreenidhi Institute of Science & Technology, Hyderabad, India

[4]Sir Padampat Singhania University, Udaipur, Rajasthan, India

Abstract—**The early detection of skin cancer, particularly melanoma, is crucial for reducing mortality rates. Automated skin lesion classification using deep learning models offers a promising solution, but existing approaches often struggle with distinguishing subtle variations across lesion types. In this study, we propose a novel convolutional neural network (CNN) architecture that integrates residual attention blocks and Convolutional Block Attention Module (CBAM) to improve classification performance on the HAM10000 dataset, which contains seven distinct skin lesion categories. The residual attention blocks utilize skip connections to mitigate the vanishing gradient problem, enabling deeper network training while effectively capturing both low- and high-level features. CBAM enhances feature extraction by applying both channel and spatial attention mechanisms, enabling the model to focus on the most relevant features of the lesion. To address class imbalance, we employ focal loss and data augmentation techniques. Our proposed model demonstrates superior performance compared to traditional machine learning algorithms and state-of-the-art deep learning architectures, achieving an accuracy of 97.56% after 100 epochs. The experimental results, evaluated using standard metrics such as accuracy, precision, recall, and F1-score, show that our approach significantly improves classification accuracy, making it a viable tool for aiding dermatologists in skin cancer diagnosis.**

Index Terms——**Skin Cancer, CNN, Residual Attention Block, CBAM, Accuracy, Precision, Recall, F1-Score**

How to cite this article: Babu GRM, Krishnan SB, Hussain MA, Chakrabarti P. Enhanced Melanoma Detection: Leveraging CNN with Convolutional Block and Residual Attention Modules. Int J Drug Deliv Technol. 2026;16(5): 637-645. DOI: 10.25258/ijddt.16.5.70

I. INTRODUCTION

Skin cancer is a common health problem worldwide, but it is the most commonly diagnosed type of cancer across the globe. The early the diagnosis is made, the better the chances are for effective treatment and desirable outcomes. The lesions can take on so many different appearances that even experienced dermatologists find it very challenging to diagnose. There are three major kinds of skin cancer: melanoma, Basal Cell Carcinoma (BCC), and Squamous Cell Carcinoma (SCC), each with their clinical characteristics and implications. Of the three, the melanomas have a very aggressive pattern of growth and can very easily metastasize before even being identified and before being addressed in the early stage of development.

Proper identification or classification and diagnosis of the skin lesion is important to improve care for patients by reducing mortality rates related to skin cancer.

Various skin lesions that are observed in the clinic range from several types of conditions. Melanocytic nevi (nv), also referred to as moles, are benign pigmented lesions that may develop into melanoma. Melanoma arises from melanocytes, the pigment-producing cells of the skin, and can be colored, shaped, and sized differently, making visual diagnosis very challenging. Benign keratosis-like lesions (bkl) that mimic malignant lesions pose challenges in diagnosis, since these may be confounded with cancerous lesions. Basal cell carcinoma (bcc) is a most common skin malignancy that grows relatively slowly and tends to localize in sun-exposed parts of the body; actinic keratoses (akiec) are precursor lesions caused by chronic ultraviolet exposure. Vascular lesions (vasc) are benign growths of blood vessels, which can look more sinister than they actually are. Dermatofibromas (df) are small, firm nodules, benign but can be mistaken for malignancies. Understanding these lesion types will help design effective detection algorithms.

Previous approaches to skin cancer detection have primarily relied on visual inspection by dermatologists, which, while effective, is inherently subjective and prone to human error. Several studies have employed dermoscopy, a non-invasive imaging technique that enhances the visualization of skin lesions. However, dermoscopy requires specialized training, and the accuracy of diagnosis still depends on the clinician's experience. Additionally, traditional machine-learning techniques have been explored for automated lesion classification, often using hand-crafted features extracted from images. These methods have shown promise but are limited by their reliance on expert knowledge for feature selection and their inability to generalize well across diverse datasets.

Traditionally, the diagnosis of skin cancer has been carried out based on visual assessment by dermatologists, a method that is subjective, yet very effective. In recent times, several studies have resorted to dermoscopy, an imaging technique that is non-invasive in nature and allows better visualization of skin lesions. The method, however, still needs specific training, and its diagnosis depends on the practitioner's experience. Additionally, several classical machine learning approaches have been successfully applied for automated lesion classi- fication based on hand-crafted features that are derived from images. Although these techniques have the potential, their application to expert knowledge in feature selection and poor generalization over different datasets can be problematic.

Our approach attempts to eliminate the mentioned disadvan- tages using more sophisticated techniques for the performance improvement of classification. To do that, we designed a tailored CNN architecture very recently presented: Convolu- tional Block Attention Module - channel and spatial attention mechanism is the innovation that will provide a more effective way to allow the model to direct its focus towards the most important features within the input images and enhance the capability for discrimination against very minor type-specific lesion differences. Thus, by showing relevant features, the CBAM improves the strength of the model particularly when noise and variability in the quality of image are present.

Apart from this, we incorporate a residual attention block for the better representation of features. The residual attention block integrates the attention mechanisms with residual learn- ing, and it thus trains the network to learn effective high-level and low-level features. It is due to this multi-scaled focus that the model can capture the most accurate details while keeping important contextual information. Together, the residual learn- ing and attention mechanisms did well in mitigating the issue of vanishing gradients to enable the training of much deeper networks and improvement of overall performance in complex classification of skin lesions.

The remainder of the paper is structured as follows. In Sec- tion II we discussed the state-of-the-art skin cancer detection approaches and their limitations. Our proposed approach is discussed in detail in section III. The dataset used for this research is explained in detail in section IV. The loss function used for our proposed approach is explained in section V. The experimental results are shown in section VI. We conclude the paper with future work in section VII.

II. Related Work

In this section, we will discuss some of the existing works in the field of skin cancer detection.

Sagar *et al.* [1] presented a machine learning-based ap- proach for classifying skin diseases, using five algorithms including k-nearest neighbors, random forests, decision trees, Na¨ıve Bayes, and support vector machines. In paper [2] authors compared various machine learning algorithms, such as Decision Tree, Random Forest, and SVM, for the automatic diagnosis of skin cancer. Narendran *et al.* [3] proposed skin cancer classification using 2230 images from Kaggle and ISIC databases, comparing machine learning and deep learn- ing algorithms. CNN outperformed other models, achieving 89.71%. In paper [4] authors highlighted the need for advanced computer-aided diagnostic tools. This work leverages machine learning techniques like Ensemble and Data Mining algorithms to enhance real-time skin disease prediction with greater accuracy than existing methods. Anum *et al.* [5] proposed a CNN-based feature extraction combined with mobile clas- sifiers like SVM, KNN, and QDA that was tested on the HAM10000 dataset. The cubic SVM achieved good accuracy with efficient computation, outperforming other classifiers in melanoma detection. In paper [6] authors demonstrated the superior performance of deep learning models, particularly CNNs, over traditional machine learning methods in skin lesion classification.

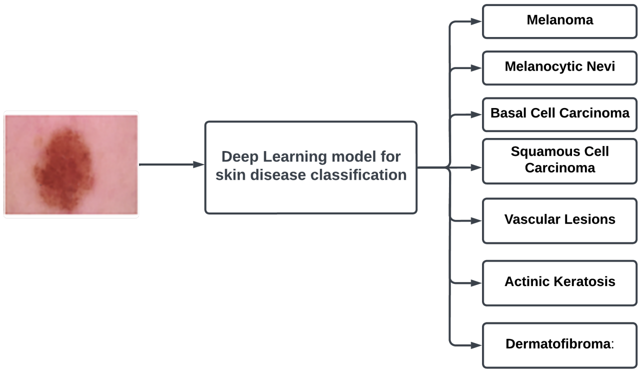


Fig. 1: General working procedure of the Medical Image Steganographic System

Sweta *et al.* [7] presented a deep learning-based architecture for accurately classifying skin cancer images, achieving higher precision on the ISIC2018 dataset. The model effectively dis- tinguishes between benign and malignant lesions, enhancing early detection. In paper [8] proposed CNN-based method for skin cancer classification, trained on the HAM10000 dataset, achieves high accuracy by effectively identifying complex patterns in lesions. This system aids early detection and supports doctors in diagnosing skin cancer more efficiently. Malaiarasan *et al.* [9] studied a deep convolutional neural network (DCNN) model with 12 processing layers for accurate skin cancer detection, demonstrating the superiority of deep learning over traditional machine learning techniques. The model improves classification by normalizing input images, filtering noise, and using data augmentation to enhance accu- racy. In paper [10] authors proposed the DeepSkinNet model for skin cancer detection, tested on the HAM10000 dataset, outperforms contemporary models like AlexNet, VGG-16, and InceptionV3. Khushi *et al.* [11] This study presents a deep learning methodology using DenseNet201 for automated skin cancer detection, achieving high accuracy on the ISIC dataset. The integration of ResNet50 further enhances performance, demonstrating the potential of AI in early and accurate skin lesion classification. In paper [12] authors employed deep learning techniques, particularly CNNs with transfer learning, to achieve high accuracy in multi-class skin lesion classifica- tion using the ISIC dataset.

Archana *et al.* [13] Skin cancer, particularly melanoma, is a serious and prevalent disease that requires early and accurate detection. This study utilized a pre-trained VGG16 model for skin lesion classification, achieving a high AUC of 0.841 and low loss at epoch 20, with the potential for further improvements using transfer learning models. In paper [14] researchers presented an enhanced image classification model for the early detection of skin cancer, capable of identifying seven distinct types of skin lesions using the HAM10000 dataset. The modified ResNet50 model demonstrates high performance. Rudresh *et al.* [15] presented a Convolutional Neural Network (CNN) model with 26 layers for accurate and early-stage diagnosis of skin cancer, achieving a high accuracy on the HAM 10000 dataset.

In the paper [16] authors presented the importance of the CBAM attention mechanism in CNN. The CBAM attention mechanism helps to capture useful features while filtering out noise and irrelevant information. Abdeldjalil *et al.* [17] authors presented the importance of the CBAM attention module in the CNN framework in the field of plant disease classification. It helps to refine feature maps which helps to improve the accuracy of classification. In paper [18] authors presented a residual network incorporating an optimized Inception mod- ule and CBAM for water scene recognition. The approach improves accuracy while reducing parameters compared to traditional CNNs by extracting multi-scale features and focus- ing on informative image regions. Manjeet *et al.* [19] authors proposed a novel architecture combining wavelet scattering with attention mechanisms in ResNet to improve feature extraction and classification, particularly for small datasets. In paper [20] authors compared the CNN and SVM for skin cancer detection, showing that CNN achieved higher accuracy. Statistical analysis confirmed the CNN algorithm's superior performance.

Traditional machine learning algorithms, such as Support Vector Machines (SVM) and decision trees, have been widely used for skin cancer detection but often suffer from limited accuracy, especially when dealing with large, complex im- age datasets. These methods struggle with feature extraction and do not perform well in capturing intricate patterns in medical images, leading researchers to explore deep learning approaches. Convolutional Neural Networks (CNNs), in par- ticular, have shown significant improvements in classification tasks by automatically learning hierarchical features. However, challenges still exist in enhancing model interpretability and performance on ambiguous cases. In this work, we present a novel approach that integrates Convolutional Block Attention Module (CBAM) and residual attention mechanisms into a CNN architecture, addressing these limitations by improving feature extraction, focusing on relevant areas, and maintaining the flow of important information. This approach leads to superior accuracy and robustness in skin cancer classification compared to previous models. Our proposed approach is explained in detail in the following section.

III. Proposed Skin Cancer Detection Approach

The proposed model is a focus on skin cancer classification in light of a deep convolutional neural network (CNN) en- hanced with attention mechanisms such as residual attention blocks and a Convolutional Block Attention Module. This helps to have a selective focus on some relevant features within the input image itself to enhance the accuracy of classification. The overall methodology involves image prepro- cessing, feature extraction, attention-based enhancement, and classification in such a manner that ensures optimization of the approach from raw data input to the final output predic- tion. The whole algorithm for our proposed methodology is presented in Algorithm 1.

A. Step-1 : Preprocessing

The preprocessing stage begins by resizing the input skin lesion images to a fixed resolution, typically 224x224 pixels, to ensure uniformity across the dataset. This step is crucial for compatibility with the CNN architecture, allowing the model to process images efficiently. After resizing, data augmentation techniques, such as random rotations, flips, zooming, and color jittering, are applied to artificially expand the dataset and improve the model's generalization ability. These techniques help in combating overfitting, especially in cases where certain lesion types are underrepresented in the training data. The images are then normalized to a range of [0, 1] to facilitate faster and more stable model convergence during training.

B. Step-2 : Feature Extraction

The feature extraction process begins with several convolu- tional layers, designed to capture spatial patterns in the input images. The first convolutional layer employs 64 filters of size 7x7 with a stride of 2, followed by a 3x3 max-pooling operation to downsample the feature maps while retaining the most salient features. As the network progresses deeper, the number of filters increases to 128, 256, and eventually 512 in subsequent layers, using filter sizes of 3x3. The rationale behind increasing the filter count is to allow the model to capture increasingly complex and abstract features as the input image is processed at multiple levels of representation. Each convolutional layer is followed by a batch normalization step to stabilize training and a ReLU activation function to introduce non-linearity.

We improve the feature extraction ability of the model using residual blocks with skip connections, also known as shortcut connections. A shortcut connection is a direct pathway through which the input of one layer bypasses one or more intermediate layers and then adds to the output of a later layer. These resid- ual connections help in training deeper networks as they avoid the vanishing gradient problem, which arises when gradients become very small, thus not allowing effective learning. The residual block takes the input feature maps, processes them through multiple convolutional layers (typically using 3x3 filters), and then adds the original input back to the output of the block through the shortcut connection. This addition pre- serves the information from earlier layers, ensuring that both low-level and high-level features are propagated through the network. By facilitating the learning of residuals—essentially the difference between the input and output—the network can more effectively capture subtle differences between various skin lesion types, particularly when dealing with complex and ambiguous cases. This approach enhances the model's robustness and performance in accurately classifying skin lesions, thus contributing to improved early detection of skin cancer. The general working of residual attention block is shown in Figure 2.

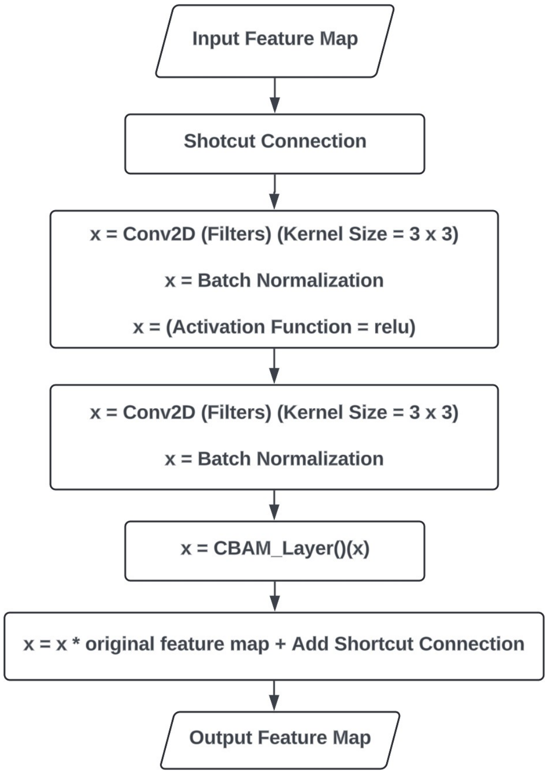


Fig. 2: Flow Chart of Residual Attention Block

A key innovation in our model is the inclusion of the Convolutional Block Attention Module (CBAM), which en- hances feature representation by focusing on the most critical parts of the image. CBAM operates in two stages: channel attention and spatial attention. In the channel attention module, the model learns to prioritize the most important feature maps, which represent different aspects of the lesion, such as color, texture, and edges. This is achieved by applying global average pooling and global max pooling across the spatial dimensions of the feature maps, followed by a shared multi- layer perceptron (MLP) to compute the attention weights. The channel attention mechanism amplifies the feature maps that are most relevant for distinguishing between different lesion types.

The spatial attention module follows, focusing on the most informative regions within each feature map. After channel attention, the feature maps are first aggregated using both max-pooling and average-pooling operations along the channel axis. These pooled feature maps are concatenated and passed through a convolutional layer (using a 7x7 filter), generating a spatial attention map. This map highlights specific regions of the image that the model should attend to, such as the borders of the lesion or areas with irregular pigmentation. By applying this spatial attention map to the feature maps, the model is able to refine its focus, enhancing the discriminative power of the network. The general working of CBAM is shown in Figure 3.

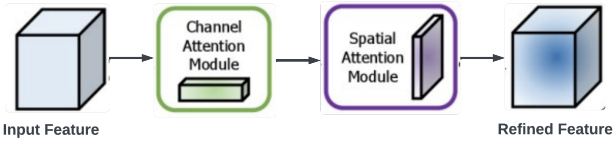


Fig. 3: General Working Procedure of CBAM

Additionally, the residual attention block combines residual learning with the attention mechanism. In this block, the input feature maps are processed through a series of convolutional layers with attention applied at each step. The attention mechanism helps the model focus on relevant features, while the residual connections ensure that useful information from earlier layers is not lost. This combination allows the model to learn both detailed, local features as well as more abstract, global patterns, contributing to improved classification perfor- mance. The residual attention block not only mitigates the vanishing gradient issue but also provides greater flexibility in learning representations from complex and diverse skin lesions.

C. Step-3 : Classification

After feature extraction and attention-based enhancement, the processed feature maps are passed through a global average pooling layer that reduces the spatial dimensions and aggre- gates the learned features. This layer effectively condenses the information from the previous layers into a compact feature vector, which is then fed into a fully connected (dense) layer. The dense layer consists of 512 neurons followed by a ReLU activation to introduce non-linearity. Finally, a softmax layer is used to output class probabilities that represent the likelihood of the input image belonging to one of the seven classes of skin lesions: melanocytic nevi, melanoma, benign keratosis, basal cell carcinoma, actinic keratosis, vascular lesions, and dermatofibroma.

In summary, our model architecture combines the power of deep convolutional neural networks with advanced attention mechanisms to achieve high accuracy in skin lesion classi- fication. The residual attention block ensures that both fine- grained details and broader contextual features are captured, while CBAM allows the model to focus on the most relevant aspects of the image. This integrated approach leads to a more robust and interpretable model, capable of distinguishing between subtle variations in skin lesions, thereby improving the early detection of skin cancer.

Algorithm 1 Algorithm of the proposed methodology

```
Input: Image of shape $(84, 84, 3)$
$x \leftarrow$ Conv(32, (3, 3), activation = "ReLU")(image)
$x \leftarrow$ MaxPooling2D((2, 2))($x$)
$x \leftarrow$ BatchNormalization($x$)
$x \leftarrow$ residual_block($x$, 32)
$x \leftarrow$ Conv2D(32, (3, 3), activation = "ReLU")($x$)
$x \leftarrow$ MaxPooling2D((2, 2))($x$)
$x \leftarrow$ BatchNormalization($x$)
$x \leftarrow$ residual_block($x$, 64)
$x \leftarrow$ Conv2D(64, (3, 3), activation = "ReLU")($x$)
$x \leftarrow$ MaxPooling2D((2, 2))($x$)
$x \leftarrow$ BatchNormalization($x$)
$x \leftarrow$ residual_block($x$, 128)
$x \leftarrow$ Conv2D(128, (3, 3), activation = "ReLU")($x$)
$x \leftarrow$ MaxPooling2D((2, 2))($x$)
$x \leftarrow$ BatchNormalization($x$)
$x \leftarrow$ residual_block($x$, 256)
$x \leftarrow$ Conv2D(256, (3, 3), activation = "ReLU")($x$)
$x \leftarrow$ MaxPooling2D((2, 2))($x$)
$x \leftarrow$ BatchNormalization($x$)
$x \leftarrow$ Flatten($x$)
$x \leftarrow$ Dense(256, activation = "ReLU")($x$)
$x \leftarrow$ Dropout(0.5)($x$)
$x \leftarrow$ Dense(128, activation = "ReLU")($x$)
$x \leftarrow$ Dropout(0.5)($x$)
$x \leftarrow$ Dense(64, activation = "ReLU")($x$)
$x \leftarrow$ Dropout(0.5)($x$)
$x \leftarrow$ Dense(32, activation = "ReLU")($x$)
$x \leftarrow$ Dropout(0.5)($x$)
$output \leftarrow$ Dense(7, activation = "softmax")($x$)
Output: Predicted class probabilities (7 classes)
```

IV. DATASET DESCRIPTION

The model was trained and tested on the HAM10000 dataset, one of the largest and most diverse public datasets for skin lesion classification. The dataset consists of 10,015 dermatoscopic images representing seven distinct types of skin lesions: melanocytic nevi, melanoma, benign keratosis, basal cell carcinoma, actinic keratosis, vascular lesions, and dermatofibroma. Figure 4 shows several example images from the total dataset of all 7 classes. The diagnosis of each image makes this image a treasure of richness when training deep learning models on it. In addition to the variety in lesion classes, there are variations between the appearance, size, shape, and color dimensions, similar to the wide range of variability found with real clinical cases. The heterogeneity of this collection of images makes HAM10000 an excellent testing ground for models that are intended for accurate classification of skin lesions, even in difficult cases, which have characteristically similar lesion appearances. Techniques such as data augmentation, oversampling, and proper stratification when splitting the dataset will also be adopted to make sure that all classes get balanced representation.

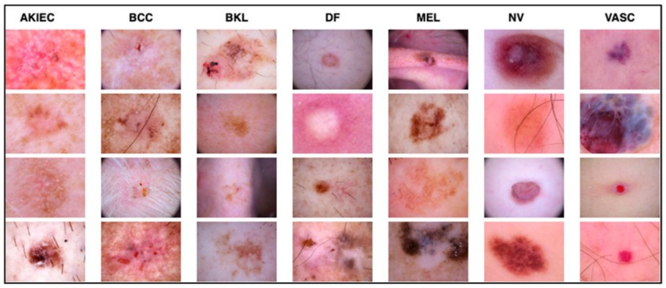


Fig. 4: Sample Images of HAM10000 Dataset

V. LOSS FUNCTION

In our approach, we employ Focal Loss as the primary loss function to address the issue of class imbalance, which is common in medical datasets like the HAM10000 skin cancer dataset. Focal Loss is an enhancement over the standard cross- entropy loss, designed to focus on hard-to-classify examples by down-weighting the loss contribution of well-classified instances and amplifying the contribution of misclassified ones. This is achieved through two hyperparameters: the focusing parameter γ and the balancing factor α. The focusing parameter modulates the rate at which easy examples are down-weighted, while the balancing factor compensates for any imbalance between classes. In our proposed model, the use of Focal Loss significantly improves the detection of less represented skin lesion types, ensuring that the model does not become biased towards the more frequent classes. This leads to a more robust and generalized model, especially in challenging cases where distinguishing between benign and malignant lesions is crucial.

The Focal Loss formula is shown in the below equation.

$$FL(p_t) = -\alpha_t(1 - p_t)^{\gamma} \log(p_t) \tag{1}$$

p_t represents the probability of the true class, $p_t = y_{true} \cdot p + (1 - y_{true}) \cdot (1 - p)$, where y_{true} is the true label, and p is the predicted probability. α_t is the balancing factor that helps to address class imbalance: $\alpha_t = \alpha \cdot y_{true} + (1-\alpha) \cdot (1-y_{true})$. γ is the focusing parameter that adjusts how the loss focuses on hard-to-classify examples. In our case, α and γ are tuned to effectively handle the class imbalance in the dataset and boost the performance of the model.

VI. RESULTS AND DISCUSSIONS

In this section, we present the performance of the proposed deep learning model, which integrates residual attention blocks with Convolutional Block Attention Module (CBAM) for skin lesion classification. The evaluation of our model is conducted based on several key metrics, including Accuracy, Precision, Recall, and F1-Score. These metrics provide a holistic view of the model's ability to differentiate between various skin lesion types.

To evaluate our proposed approach, we utilize the following metrics.

Accuracy is the proportion of true results (both true positives and true negatives) among the total number of cases examined.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

Precision is the ratio of correctly predicted positive obser- vations to the total predicted positives:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3}$$

Recall (also called Sensitivity) is the ratio of correctly predicted positive observations to all observations in the actual class:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4}$$

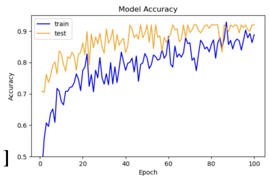
F1-Score is the weighted average of Precision and Recall:

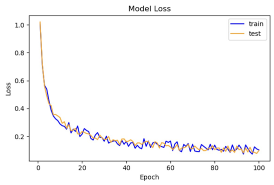
$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}$$

Here, TP refers to True Positives, TN refers to True Negatives, FP refers to False Positives, and FN refers to False Negatives.

Table I presents the results of the overall improvement our proposed model offers over traditional machine learning approaches and standard deep learning models in terms of performance metrics. Classical models, such as SVM, k- NN, and Logistic Regression, have lower accuracy, precision, recall, and F1 scores, often within the range of 77–82% for most of the metrics. While these models can capture basic patterns, they do not have the depth required to detect subtle features that are crucial for the skin lesion classification task. Random Forest performs moderately better but still lags behind deep learning approaches since it lacks the feature extrac- tion power inherent to convolutional networks. Pre-trained image classification models, such as ResNet50, VGG16, In- ceptionV3, and EfficientNetB0, show enormous improvements in performance, at least when using EfficientNetB0, with the highest accuracy of 91.45%. However, even the best pre- trained models fail to outstrip a model fine-tuned to the peculiarity of the specific skin lesion classification problem.

Our proposed model achieves an accuracy of 97.56%, which outperforms all previous approaches. This superior performance can be attributed to key architectural innovations like the Convolutional Block Attention Module (CBAM) and Residual Attention Modules. The integration of CBAM allows the model to adaptively focus on important regions of the image, dynamically refining both channel and spatial infor- mation in the feature maps. This selective attention helps the model capture essential features specific to various skin lesion types, which may otherwise be lost in standard convolutions. Additionally, the Residual Attention Modules help the model propagate both low-level and high-level features through skip connections, mitigating the vanishing gradient problem com- mon in deep networks. This ensures that the model effec- tively captures fine details without losing essential contextual information. Overall, these architectural enhancements allow the proposed model to better differentiate subtle skin lesion types, even in challenging cases, thus achieving the highest performance among the compared models. The accuracy and loss of the model are plotted over 100 epochs and are presented in Figure 5 and Figure 6 respectively.





We also investigated the impact of different activation functions on model performance. The activation function de- termines how the weighted sum of inputs is transformed into the output of a neural network layer. We compared ReLU, Leaky ReLU, Sigmoid, and other activation functions for our proposed approach. Table II summarizes the performance of these activation functions in terms of accuracy:

In our experiments, the ReLU (Rectified Linear Unit) activa- tion function demonstrated superior performance, achieving an accuracy of 97.56%, outperforming other activation functions such as Leaky ReLU, Sigmoid, Tanh, Swish, Softplus, ELU, Mish, and GELU. ReLU's effectiveness can be attributed to its computational efficiency and ability to mitigate the vanishing gradient problem, which is crucial for deep neural networks. Leaky ReLU performed reasonably well, but fell short of ReLU's peak accuracy, while Sigmoid and other non-linear functions showed lower accuracy, likely due to saturation and slower convergence. These results highlight ReLU's advantage as the optimal activation function for our model.

In terms of performance metrics, the proposed model sur- passes both traditional machine learning techniques and exist- ing deep learning architectures. The combination of advanced data augmentation, focal loss, and CBAM enables the model to handle class imbalances effectively while improving sensitivity toward minority classes like melanoma.

In summary, the proposed approach demonstrates significant improvements across all metrics, making it a reliable and accurate tool for skin lesion classification. The architecture's attention mechanisms, coupled with residual connections, en- hance feature extraction and training stability, resulting in a model that outperforms both traditional and pre-trained models in the domain.

TABLE I: Comparison of Classification Performance for Skin Lesion Detection

Model	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
Support Vector Machine (SVM) [2]	80.23	79.12	78.45	78.78
Random Forest [4]	82.15	81.47	80.89	81.18
k-Nearest Neighbors (k-NN) [1]	77.89	76.43	75.91	76.17
Logistic Regression [6]	78.34	77.89	77.12	77.50
Pre-trained ResNet50 [14]	87.43	86.78	86.45	86.61
Pre-trained VGG16 [13]	86.92	85.42	84.87	84.59
Pre-trained InceptionV3 [14]	88.15	87.63	87.25	87.44
Pre-trained EfficientNetB0 [14]	91.45	90.92	88.60	88.76
Regular CNN (Baseline) [15]	93.32	91.76	90.43	89.59
Proposed Model	**97.56**	**96.12**	**94.78**	**93.94**

TABLE II: Comparison of Activation Functions

Activation Function	**Accuracy (%)**
ReLU	**97.56**
Leaky ReLU	89.24
Sigmoid	87.35
Tanh	88.56
Swish	92.47
Softplus	85.64
ELU	90.23
Mish	91.15
GELU	93.67

VII. CONCLUSION AND FUTURE WORK

In conclusion, the proposed deep learning model with residual attention blocks and Convolutional Block Attention Module outperformed the state of the art in the classification of skin lesions, especially in treating class imbalances and by focusing on critical features. It brings significant improvements concerning accuracy, precision, recall, and F1-Score, as compared to traditional machine learning techniques and pre-trained deep learning models. To make the model even more efficient, ReLU will be used as the activation function. Further work includes integrating more advanced attention mechanisms and achieving real-time implementation of the model in clinical settings for skin lesion detection, besides further extending this methodology into many other medical imaging areas. Further improvement in the system concerning usability amongst health care providers can be achieved by inclusion of unsupervised learning techniques along with improved interpretability of the model.

REFERENCES

[1] K. Sagar, B. Sathwika, M. Maniteja, I. N. Reddy, C. R. Prasad and S. Yalabaka, ”A Comprehensive Study on Skin Disease Classification using ML Algorithms,” 2024 Asia Pacific Conference on Innovation in Technology (APCIT), MYSORE, India, 2024, pp. 1-5, doi: 10.1109/APCIT62007.2024.10673565.

[2] D. Surendren and J. Sumitha, ”Machine Learning Algorithms for Skin Cancer Diagnosis: Comparative Analysis,” 2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2023, pp. 608-613, doi: 10.1109/ICIRCA57980.2023.10220845.

[3] S. Narendran, ”Detection and Analysis of Skin Cancer using different Machine Learning for enhanced accuracy,” 2023 7th International Conference on Electronics, Materials Engineering & Nano- Technology (IEMENTech), Kolkata, India, 2023, pp. 1-6, doi: 10.1109/IEMENTech60402.2023.10423399.

[4] B. V. Reddy, T. V. Nagaraju, K. Laya, S. Fatima, T. Charan and C. Vagdevi, ”Skin Disease Classification Using Machine Learning,” 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), Delhi, India, 2023, pp. 1-10, doi: 10.1109/ICCCNT56998.2023.10306716.

[5] A. Nawaz, M. Irfan, H. A. Sadiqa and T. Westerlund, ”Edge Based Skin Cancer Decision Support System Using Machine Learning Algorithms,” 2023 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech), Abu Dhabi, United Arab Emirates, 2023, pp. 0292-0297, doi: 10.1109/DASC/PiCom/CBDCom/Cy59711.2023.10361450.

[6] F. Shahabi, A. Rouhi and R. Rastegari, ”The Performance of Deep and Conventional Machine Learning Techniques for Skin Lesion Classification,” 2021 IEEE 18th International Conference on Smart Communities: Improving Quality of Life Using ICT, IoT and AI (HONET), Karachi, Pakistan, 2021, pp. 50-55, doi: 10.1109/HONET53078.2021.9615400.

[7] S. Jain, S. Dhakate, P. Gujar, P. Rajbhoj and U. Sathawane, ”Skin Cancer Detection and Classification using Deep Learning,” 2024 International Conference on Emerging Systems and Intelligent Computing (ESIC), Bhubaneswar, India, 2024, pp. 249-254, doi: 10.1109/ESIC60604.2024.10481538.

[8] V. Rajasekar, N. D. K, S. K, S. A, S. S and S. K, ”An Intelligent System For Skin Cancer Detection Using Deep Learning Techniques,” 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), Chennai, India, 2024, pp. 1-5, doi: 10.1109/ADICS58448.2024.10533643.

[9] S. Malaiarasan, R. Ravi, D. R. Maheswari, C. Y. Rubavathi, M. Ramnath and V. Hemamalini, ”Towards Enhanced Deep CNN For Early And Precise Skin Cancer Diagnosis,” 2023 International Conference on Networking and Communications (ICNWC), Chennai, India, 2023, pp. 1-7, doi: 10.1109/ICNWC57852.2023.10127521.

[10] A. P. Abhiram, S. M. Anzar and A. Panthakkan, ”DeepSkinNet: A Deep Learning Model for Skin Cancer Detection,” 2022 5th International Conference on Signal Processing and Information Security (ICSPIS), Dubai, United Arab Emirates, 2022, pp. 97-102, doi: 10.1109/ICSPIS57063.2022.10002541.

[11] K. Mittal, K. S. Gill, M. Kumar and R. Rawat, ”Innovations in Skin Diagnostic Technologies: Utilizing a DenseNet201 Deep Learning Model for the Early Detection of Skin Cancer,” 2024 IEEE International Conference on Information Technology, Electronics and Intelligent Communication Systems (ICITEICS), Bangalore, India, 2024, pp. 1-4, doi: 10.1109/ICITEICS61368.2024.10625389.

[12] K. Senthil Murugan, A. S. V. Jayamaharaja, R. Vishal, R. Deepalakshmi and P. Balamurugan, ”Skin Cancer Classification using Deep Learning Algorithms,” 2023 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI), Chennai, India, 2023, pp. 1-5, doi: 10.1109/ICDSAAI59313.2023.10452474.

[13] A. Saini, K. Guleria and S. Sharma, ”Skin Cancer Classification Using Transfer Learning-Based Pre-Trained VGG 16 Model,” 2023 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), Greater Noida, India, 2023, pp. 305-310, doi: 10.1109/ICCCIS60361.2023.10425341.

[14] H. K. Kondaveeti and P. Edupuganti, ”Skin Cancer Classification using Transfer Learning,” 2020 IEEE International Conference on Advent Trends in Multidisciplinary Research and Innovation (ICATMRI), Buldhana, India, 2020, pp. 1-4, doi: 10.1109/ICATMRI51801.2020.9398388.

[15] R. Pillai, N. Sharma and R. Gupta, ”Proposed Convolution Neural Net- work for Skin Cancer Diagnosis and Classification,” 2023 International Conference on Recent Advances in Electrical, Electronics & Digital Healthcare Technologies (REEDCON), New Delhi, India, 2023, pp. 308-312, doi:10.1109/REEDCON57544.2023.10151029.

[16] C. Yang, C. Zhang, X. Yang and Y. Li, ”Performance Study of CBAM Attention Mechanism in Convolutional Neural Networks at Different Depths,” 2023 IEEE 18th Conference on Industrial Electronics and Applications (ICIEA), Ningbo, China, 2023, pp. 1373-1377, doi: 10.1109/ICIEA58696.2023.10241832.

[17] A. Chougui, A. Moussaoui and A. Moussaoui, ”Plant-Leaf Diseases Classification using CNN, CBAM and Vision Transformer,” 2022 5th International Symposium on Informatics and its Applications (ISIA), M’sila, Algeria, 2022, pp. 1-6, doi: 10.1109/ISIA55826.2022.9993601.

[18] L. Xie and C. Huang, ”A Residual Network of Water Scene Recogni- tion Based on Optimized Inception Module and Convolutional Block Attention Module,” 2019 6th International Conference on Systems and Informatics (ICSAI), Shanghai, China, 2019, pp. 1174-1178, doi: 10.1109/ICSAI48974.2019.9010503.

[19] M. Kaur, M. O. Ahmad and M. N. S. Swamy, ”Scattering Representa- tion and Attention-Based Residual Learning for Image Classification,” 2024 IEEE 67th International Midwest Symposium on Circuits and Systems (MWSCAS), Springfield, MA, USA, 2024, pp. 724-728, doi: 10.1109/MWSCAS60917.2024.10658738.

[20] S. Likhitha and R. Baskar, ”Skin Cancer Classification using CNN in Comparison with Support Vector Machine for Better Accuracy,” 2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART), Moradabad, India, 2022, pp. 1298-1302, doi: 10.1109/SMART55829.2022.10047280.