Research Article

Skin Cancer Cell Detection using Image Processing

Taskin Sabit[1*], Faiza Tasnim[2], Sadia Afrin Sara[2], Sharia Tasnim Adrita[2], Maisha Tarannum[3]

[1] *Department of Pharmacology and Toxicology, Master of Science, Wright State University, Dayton, OH, USA.*

[2] *Computer Science and Engineering, Bachelor of Science, American International University-Bangladesh, Dhaka, Bangladesh.*

[3] *Biomedical Engineering, Bachelor of Science, Bangladesh University of Engineering and Technology, Dhaka, Bangladesh.*

Article Info

Article history

Received: 25/03/2025
Revised 1: 30/04/2025
Accepted: 25/05/2025

Keywords:
Deep learning
Convolutional neural networks (CNN)
VGG16
Classification

ABSTRACT

Early diagnosis and precise detection of skin cancer represent a global health priority since this disease remains highly dangerous while being among the most frequent ones. This research investigates the effectiveness of deep learning techniques, specifically Convolutional Neural Networks (CNN) and the VGG16 architecture, for skin cancer detection and classification. The study works with images from the International Skin Imaging Collaboration (ISIC) while employing resizing and augmentation preprocessing to boost its model performance. We evaluate the proposed model using precision, recall, and F1-score metrics to ensure accurate classification. The proposed CNN model achieved 87% validation accuracy, outperforming the VGG16 model, which attained 65% accuracy. Experimental results highlight the potential of AI-driven models in improving diagnostic accuracy, demonstrating their significance in medical image analysis and early skin cancer detection.

1. Introduction

Skin cancer can occur very often and can be life-threatening, so making sure it is detected early is of great importance worldwide. Currently, some diagnostic methods depend on a physician's experience and invasive tools that might not be available when needed. Skin cancer can be defined as an unusual growth of skin cells caused by exposure to ultraviolet rays from the sun. These skin cells grow abnormally and form tumors. This research shows the efficiency of deep learning models like CNN and VGG16 architecture for early skin cancer detection. It is the most prevalent form of cancer worldwide. It can extend to other organs and tissues through the bloodstream or lymphatic system and cause further harm to the body. Individuals with light hair and red spots across their skin have a higher chance of developing skin cancer. The use of the ISIC dataset and robust preprocessing and augmentation methods helps provide a range of representative images for the research models. Authors talk about how accuracy, precision, recall and F1-score assess model performance in the study, as well as point out that custom CNNs excel over VGG16. Our goal through this system is to provide doctors with an intelligent, image-based tool that improves diagnostic accuracy for use in both clinical and resource-poor areas.

CNN and VGG16 fall under the general category of deep learning models within artificial intelligence (AI), specifically used for image classification and analysis. CNN is a type of neural network particularly effective for image processing. It automatically learns spatial hierarchies of features like edges, shapes, and textures through convolutional, pooling, and fully connected layers. CNNs are efficient in skin cancer detection because they can automatically extract features from images, are scalable and adaptable to various image tasks, and perform well with large datasets and high-dimensional image data. VGG16 is a specific CNN architecture introduced by the Visual Geometry Group at Oxford. It consists of 16 layers (13 convolutional and 3 fully connected layers) and is known for its uniform structure, using small 3x3 filters throughout. It is widely used for benchmarking in image classification tasks. VGG16 provides a well-tested, stable architecture, has proven performance in many image recognition challenges like ImageNet, and offers a good baseline for comparison and transfer learning. Comparing CNN and VGG16 helps determine which model provides better accuracy and generalization for skin cancer cell detection. CNNs can be custom-built and lighter, while VGG16 is deeper but more resource-intensive. How architecture depth and feature extraction capabilities affect diagnostic outcomes.

Three principal skin cancer forms exist: basal, squamous, and melanoma. Various kind of skin cancers exists globally. Actinic keratosis (AK) lesions typically appear as petite, rough, and crusty patches on the skin. It, also known as solar keratosis, is a scaly or crusty skin growth indicative of sustained sun damage. It may progress to invasive neoplasms and has been interpreted as the earliest sign of skin cancer [1]. Melanoma is the deadliest in comparison to other skin cancers. Skin cancer forms due to melanocyte malignant transformation and functions as one of the most lethal skin malignancies because it easily resists treatments, while causing frequent recurrence and resulting in

***Corresponding author:** Taskin Sabit
*E-mail address: taskin.sabit.wsu@gmail.com https://doi.org/10.56158/jpte.2025.122.4.01

poor patient survival. During 2014, the United States recorded 76,100 new melanoma diagnoses, and the mortality rate was estimated to be 9710 deaths. However, with a survival rate of 99%, it can be surgically removed by detecting it in an early stage [2]. This type of lesion is a relatively small circle, which varies from pink to brown while keeping either a thin tan layer or forming bigger clumps across the face, trunk, and extremities. Doctors can easily detect it. However, patients rarely receive tissue samples for examination [3]. Basal cell carcinoma (BCC) stands as the least common type of skin cancer worldwide. It is mainly caused by long-term exposure to the sun's ultraviolet rays. BCCs are slow-growing, locally invasive, epidermal skin tumors that primarily affect individuals with fair skin [4]. Cancer registries do not encompass data on BCC due to its low mortality rate in many countries. However, evaluating data from insurance registries and official statistics in the United States, BCC incidence has been estimated to reach 4.3 million cases each year, and it is common in the Caucasian population [5]. The medical description of Nevus defines it as a pigmented skin growth called a mole. Squamous cell carcinoma develops in the outer layer of skin from flat squamous cells that constitute this skin layer. head and neck region [6]. The skin condition known as dermatofibroma presents itself as a minor skin elevation called fibrous histiocytoma. People often find dermatofibroma initially on their legs, even though it can appear across any area, since this condition typically develops in female bodies. It is one of the most commonly encountered soft-tissue lesions, accounting for approximately 3% of skin lesion specimens received by dermatopathology laboratories [7]. Detecting and curing these various types of skin cancer globally is imperative.

In conclusion, this paper has demonstrated remarkable potential in skin cancer detection. It can learn intricate patterns and features from an efficient approach to identifying skin cancer variations by comparing deep-learning models. This paper utilizes and compares the accuracy of two different models: CNN and VGG16.

1.1. Related Work

Rezaoana, N., and his coauthors follow a three-phase approach in their study. Firstly, they perform data acquisition and augmentation to gather relevant information. In the model development phase, they introduce a unique parallel CNN architecture with dilated convolution layers for feature extraction and classification. Through comparative analysis against VGG-16 and VGG-19 models, the authors demonstrate the superior performance of their proposed model in terms of accuracy, precision, recall, and F1-score. This structured approach effectively showcases the ability of convolutional neural networks (CNNs) to accurately classify skin cancer while highlighting potential advancements in medical diagnostics. The authors suggest exploring alternative CNN architectures and expanding datasets to enhance the model's robustness [8]. In a related study, Rezaoana et al. also structure their approach into three main phases: gathering and enhancing data, creating the model, and making predictions. They employ various AI algorithms, including CNNs and Support Vector Machines (SVMs), and integrate these with image processing techniques to build a more effective framework. This results in an improved accuracy rate of 85% [9]. Another group of researchers proposes a comprehensive method that combines preprocessing, lesion detection, and feature extraction based on the widely accepted ABCD rule: Asymmetry, Border irregularity, Color, and Diameter, followed by classification using the Total Dermatoscopic Value (TDV). Their system achieves promising results, with an accuracy of 90% in identifying benign, suspicious, and malignant lesions. Implementing automated diagnostic systems such as these enhances the speed and accuracy of skin cancer detection. This computer-based approach enables the identification of subtle visual patterns such as asymmetry, color variation, and texture differences that may go unnoticed by the human eye, thereby reducing diagnostic time and improving the overall accuracy of clinical evaluations [10]. Haque, T. has a review that considers skin physiology, non-melanoma skin cancer (NMSC), the relationship between AK and skin cancer, and drugs administered topically for these conditions. The dermal preparations for managing NMSC and AK are discussed in detail. Notably, few studies have examined drug disposition in cancerous skin or AK. Finally, recent novel approaches for targeting drugs to skin neoplasms [1]. Naves, L. intends to elucidate the possibilities to treat melanoma skin cancer using hybrid nanofibers developed by an advanced electrospinning process. Their review shows that enhanced permeability and retention are the basis for nanotechnology, which aims at topical drug delivery. They also report a case study involving two approaches targeting melanoma skin cancer therapy: magnetic-based core–shell particles and electrospun mats [2]. Karadag, A. and his coauthors found that Seborrheic keratosis (SK), whose appearance is generally a small roundish reddish to brownish scaling lesion ranging in size from a few mm to many mm, may have a single presentation or be one of many such lesions. In recent retrospective studies and case reports, SKs have rarely been found to have malignant characteristics. Although these studies are inconclusive, lesions that are inflamed, bleeding, ulcerated, or sufficiently irritated may require being biopsied to rule out melanoma or other malignancies [3].

2. Research Methodology

This section present techniques used to detect skin cancer. The approach includes exploring and evaluating two different model architectures: CNN) and VGG16. The main objective is to compare and find out the performance of these models and identify their respective accuracy strengths and weaknesses in classifying skin cancer.

2.1. Data Collection and Preparation

2.1.1. Collecting our Data

The data used in this model is collected from the International Skin Imaging Collaboration (ISIC). Our data consists of images of 9 types of skin cancer [13]. Those are: Actinic Keratosis, Basal Cell Carcinoma, Dermatofibroma, Melanoma, Nevus, Pigmented Benign Keratosis, Seborrheic Keratosis, Squamous Cell Carcinoma, and Vascular Lesion. Among the nine different classes, pigmented benign keratosis is the most represented category, with 462 images [13]. Master images of melanoma

skin cancer represent the second most common category among 438 examples that demonstrate the importance of early diagnosis in clinical settings [13]. Humans tend to develop basal cell carcinoma as their leading skin cancer type, and the collection contains 376 images of this form [13]. The three groups of lesions under nevus encompass moles and benign pigmented lesions, which amount to 357 cases. Squamous cell carcinoma, a malignant tumor that can spread if left untreated, has 181 images in the dataset, offering significant representation for model training [13]. Less common conditions, such as vascular lesions and actinic keratosis, are included with 139 and 114 images, respectively. Dermatofibroma, a benign skin nodule, is represented by 95 samples [13]. Finally, seborrheic keratosis, a non-cancerous but often cosmetically concerning condition, is the least represented class with 77 images. This distribution of images across both malignant and benign categories provides a robust foundation for training deep learning models.

It is essential to see the balance and potential biases in the distribution of images across different classes in the dataset. To achieve this, data visualization techniques were employed. A bar plot created using the 'seaborn' library illustrates the quantity of images in each class.

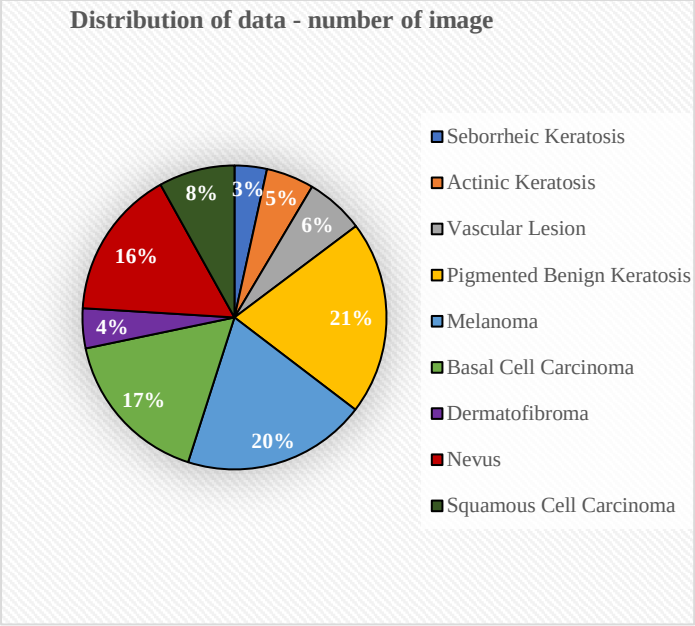


Fig. 1. Distribution of our data

The pie chart illustrates the distribution of images across different skin lesion categories in a dataset. The data consists of nine classes corresponding to specific percentages throughout the dataset. The survey analysis shows Seborrheic Keratosis being the leading identified lesion at 21%, followed by Melanoma and Basal Cell Carcinoma, together with Seborrheic Keratosis, which forms 57% of all examined skin conditions. The distribution of samples implies an adequately balanced dataset that features frequent skin lesions because it presents a wide variety of cases for medical image classification systems.

2.1.2. Data Visualization and Preprocessing

This section presents the visualization and exploration of the skin cancer dataset. It will help to provide insight into the data presented in each class. A single image was taken from each class to show the data visually and presented using the 'matplotlib' library. This will effectively provide viewers with insight into our dataset. The display arrangement arranges these images in a 3x3 grid layout, providing a visual insight into our dataset's comprehensive distribution of classes. Figure 1 represents one sample image from each type of skin cancer category.

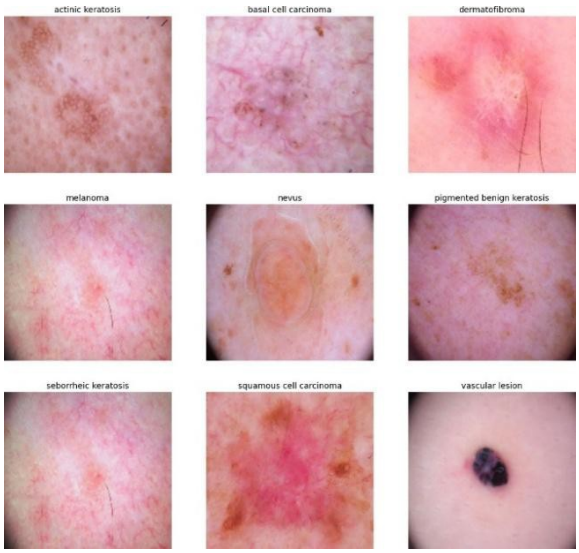


Fig. 2. Example of our data [13]

Data preprocessing is an essential part of image processing. Modifying the data is necessary to feed it into a model. Preprocessing the images ensures they are of the desired size, format, and quality. Preprocessed techniques and approaches are discussed in later sections.

2.1.3. Image Resizing for Standardization

The collected dataset is vast, and due to our limited resources, it is essential to resize our data. The original dataset downloaded from ISIC is close to 1GB, and each image is up to 15 MB. But to process those images, a powerful machine is needed. So, due to our limited resources, we resized and decreased the size of each image. This strategic decrease reduced the size of our dataset to close to 200 MB.

2.1.4. Data Augmentation for Improved Generation

The original dataset required data augmentation to tackle its class imbalance problem. The approach increased the dataset size and ensured balanced class distribution because this aspect prevents model bias and enhances generalization during training. The augmentation process was implemented using the 'Augmentor' Python library. A combination of rotation, horizontal flipping, zooming, and brightness adjustment techniques was applied to generate synthetic images from the original dataset. Rotation (±25 degrees) was primarily used, as it does not significantly alter lesion characteristics and maintains diagnostic integrity. Horizontal flipping was also employed, as

lesions can appear on either side of the body without clinical significance. Additional transformations, such as slight zooming and brightness changes, introduced variability while preserving critical features necessary for accurate classification. Each underrepresented class was augmented until it reached approximately the same number of samples as the most represented class, resulting in a more uniform dataset distribution. On average, three new samples were generated per original image for the minority classes. After augmentation, the total number of images increased from 2,239 to 6,739. This process significantly enhanced the dataset's diversity and improved the deep learning model's ability to generalize across unseen examples during testing.

Table 1. Argumentation effect in the dataset

Total images before Argumentation	Total images after Argumentation
2239	6739

2.2. Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) are a fundamental architecture in deep learning, widely used for tasks such as image recognition, object detection, and natural language processing. CNNs have shown remarkable success in various computer vision tasks, including image classification, object detection, and segmentation **[11].** CNNs are designed to effectively capture patterns in images, making them highly suitable for tasks that involve understanding complex visual data **[12].** The architectural design of CNNs is taken from the mechanism in the human brain.

2.2.1. Training Dataset Preparation

The images are organized into batches to help this model learn. Each batch contains 32 images, resized to 180 x 180 pixels. The labels (types of cancer) are also formatted into a 'categorical' format. "Caching" is used to handle the data efficiently. It loads the images into memory or RAM, helping to speed up the learning process.

2.2.2. Building CNN Model

The proposed CNN model is designed layer by layer using a sequential approach. Each layer acts as a building block, allowing it to learn specific aspects of the images. The proposed model is structured using:

Data rescaling layer: This layer rescales the pixel value of the image. It ensures all the colors are in the same range so the model can learn effectively.

Convolution layer: This layer is responsible for recognizing patterns in the image. This filter teaches us to recognize feature-like edges, texture, and shape, and passes the information to the next layer.

Pooling layer: After each convolution layer, we have a pooling layer. It zooms out and captures the most critical information from the pattern that the previous layer found. It helps reduce complexity and makes the model faster and efficient.

Dropout layer: Whenever a deep CNN model is used, that model tends to memories the training data. The dropout layer randomly selects some neurons and deactivates them. And this technique avoids overfitting.

Dense layer: This layer analyzes, combines information, and makes predictions. Our proposed architecture has two thick layers, with the first layer having 128 units and using the 'relu' activation function.

Output layer: This layer provides output. It uses the 'softmax' activation function and provides a probability of the image belonging to each type of skin cancer.

Table 2. The summary of the proposed CNN model with sequential layers, configuration, and output shape

Layer (type)	Output Shape	Param
rescaling (Rescaling)	(None, 180, 180, 3)	0
conv2d (Conv2D)	(None, 178, 178, 32)	896
max_pooling2d (MaxPooling2D)	(None, 89, 89, 32)	0
conv2d_1 (Conv2D)	(None, 87, 87, 64)	18496
max_pooling2d_1 (MaxPooling2D)	(None, 43, 43, 64)	0
conv2d_2 (Conv2D)	(None, 41, 41, 128)	73856
max_pooling2d_2 (MaxPooling2D)	(None, 20, 20, 128)	0
dropout (Dropout)	(None, 20, 20, 128)	0
flatten (Flatten)	(None, 51200)	0
dense (Dense)	(None, 128)	6553728
dropout_1 (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 9)	1161
Total params: 6,648,137 Trainable params: 6,648,137 Non-trainable params: 0		

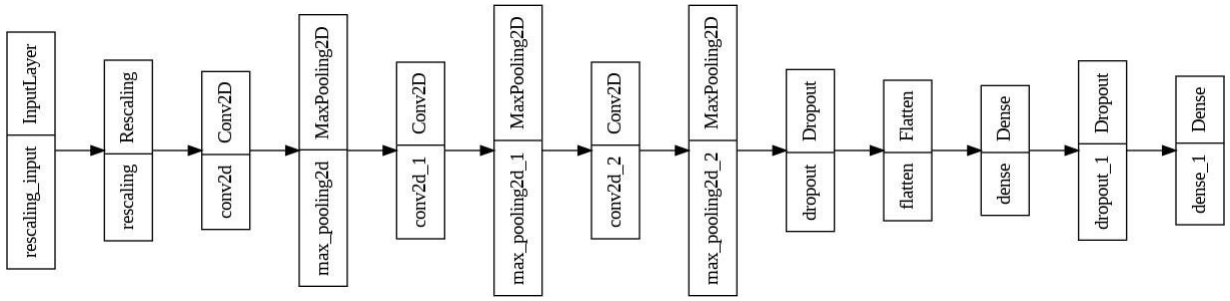


Fig. 3. A visual representation with sequential layers of the proposed model

2.2.3. Model Compilation and Training

We used the Adam optimization algorithm to train our CNN model. A variant of stochastic descent that dynamically changes the learning rate during training. The 'categorical_crossentropy' loss function was selected to model's performance in the multi-task classification task.

In this proposed model, a two-callback mechanism is used.

• Model Checkpoint

• Early Stopping

Model Checkpoint: which saves the model weight in a checkpoint file. This ensures that model progress, architecture, and learned parameters are stored in a specified interval in a file (model.h5). It monitors specified metrics, and a callback determines whether the current model has improved. If validation accuracy improves, model weights are saved.

Early Stopping: Its callbacks are designed to monitor the validation matrix (validation accuracy in our case). They stop the training process if the monitor matrix fails to show improvement for a specified number of epochs. This helps the model avoid overfitting and increases training efficiency.

The proposed model is trained on 20 epochs, representing the number of times the dataset passes through the model while training. Trading history, such as the loss curve and accuracy, is discussed in a later section.

2.3. VGG16

VGG16 is a convolutional neural network architecture, known as Visual Geometry Group 16. It contains 16 layers, including 13 convolutional and three fully connected layers. VGG16 uses small 3x3 convolutional filters and pooling layers to extract features.

2.3.1. Training Dataset Preparation

The images are organized into batches to help this model learn. Each batch contains 64 images. Images are resized to 224 x 224 pixels. The labels (types of cancer) are also formatted into a 'categorical' format.

2.3.2. Building VGG16 Model

In the 'create_model' function, the proposed architecture was built using the VGG16 architecture with an additional layer.

VGG16 Base and fine-tuning: The neural network model combines the VGG16 architecture with custom-designed layers. The VGG16 base is loaded through the VGG16 function, where the 'include_top' parameter is set to False to exclude the fully connected layers. Fine-tuning is also used, which gives the flexibility to specify a certain number of trainable or frozen layers within the VGG16 base.

Tailored layer: A custom layered is used on top of the VGG layer.

Flatten layer: This layer reshapes the output into a vector.

Dense layer: This layer analyzes, combines information, and makes predictions. Our proposed architecture has two thick layers, with the first layer having 4096 units and using the 'relu' activation function.

Dropout layer: Whenever, deep layered model is used, that model tends to memories the training data. The dropout layer randomly selects some neurons and deactivates them. And this technique avoids overfitting.

Model Compilation: After building the model architecture, the next step is to compile it. The model is compiled using the Adam optimizer, with categorical cross-entropy as the loss function and accuracy as the evaluation metric.

Table 3. The summary of the proposed VGG16 model with sequential layers, configuration, and output shape

Layer Name	Output Shape	Parameters
input_2	(None, 224, 224, 3)	0
block1_conv1	(None, 224, 224, 64)	1,792
block1_conv2	(None, 224, 224, 64)	36,928
block1_pool	(None, 112, 112, 64)	0
block2_conv1	(None, 112, 112, 128)	73,856
block2_conv2	(None, 112, 112, 128)	1,47,584
block2_pool	(None, 56, 56, 128)	0
block3_conv1	(None, 56, 56, 256)	2,95,168
block3_conv2	(None, 56, 56, 256)	5,90,080
block3_conv3	(None, 56, 56, 256)	5,90,080
block3_pool	None, 28, 28, 256)	0
block4_conv1	(None, 28, 28, 512)	11,80,160
block4_conv2	(None, 28, 28, 512)	23,59,808
block4_conv3	(None, 28, 28, 512)	23,59,808
block4_pool	(None, 14, 14, 512)	0
block5_conv1	(None, 14, 14, 512)	23,59,808
block5_conv2	(None, 14, 14, 512)	23,59,808
block5_conv3	(None, 14, 14, 512)	23,59,808
block5_pool	(None, 7, 7, 512)	0
flatten	(None, 25088)	0
dense_3	(None, 4096)	10,27,64,544
dense_4	(None, 1072)	43,91,984
dropout_1	(None, 1072)	0
dense_5	(None, 9)	9,657
Total params: 121,880,873 Trainable params: 107,166,185 Non-trainable params: 14,714,688		

Table 3 summarizes the proposed VGG16 model, with sequential layers, configuration, output shape, and parameter number in each layer.

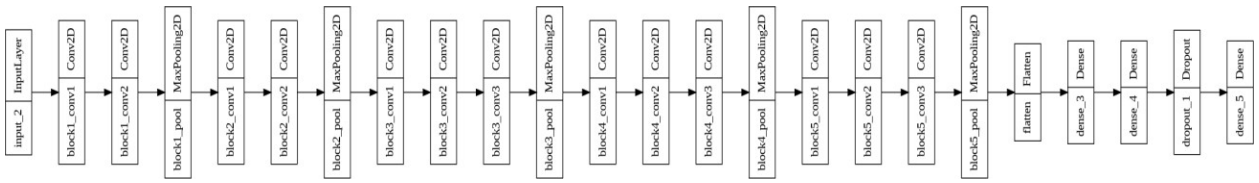


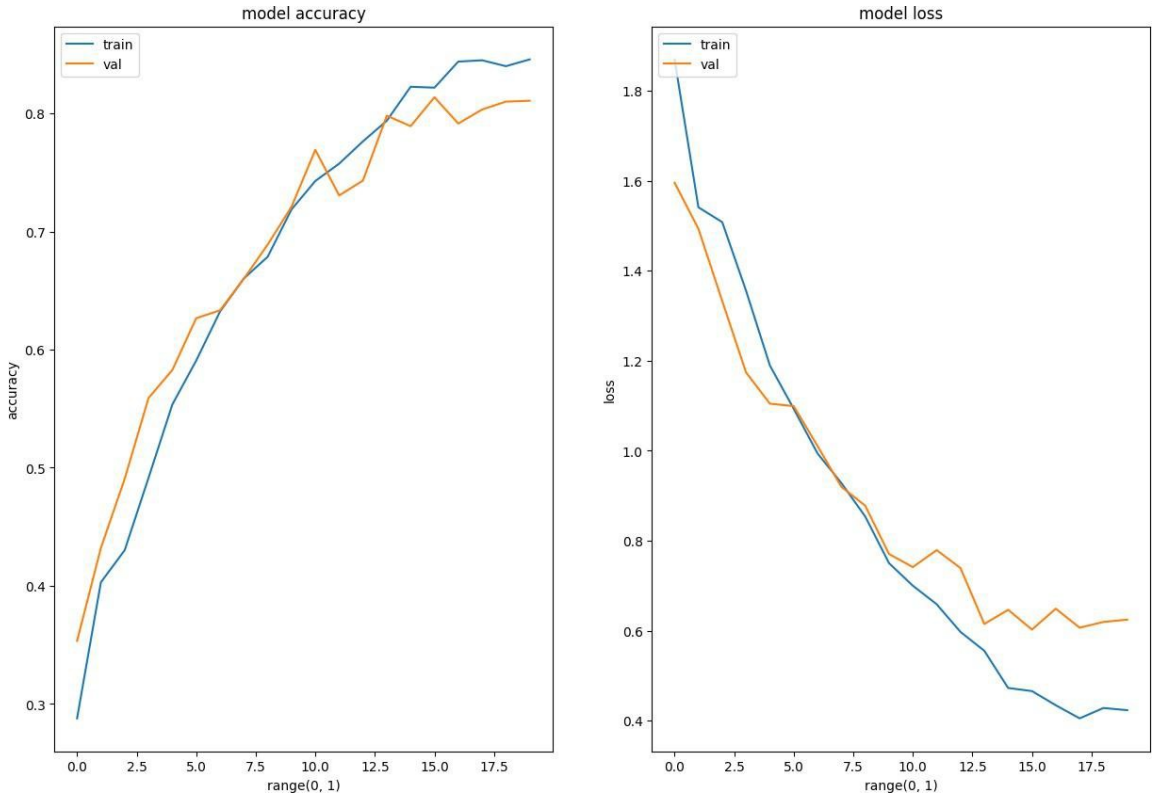
Fig. 4. A visual representation with sequential layers of the proposed VGG16 model

3. Results and Findings

3.1. Proposed CNN model result

Various analyses evaluated the proposed skin cancer detection model's performance. This section explains the key findings observed during the model's training and validation phases.

Figure 5 represents a visualization and the changes in accuracy and loss in the training phase. The x-axis represents the number of training epochs, and the y-axis represents the corresponding accuracy or loss value. The accuracy curve in Figure 5 (a) shows the model's ability to classify skin cancer images correctly. This graph shows that the training and validation accuracy increase while the number of epochs increases. By observing this graph, shows this model has no overfitting because the training and validation curves are close to each other. Figure 5 (a) shows that this model can classify with 87% accuracy. The loss curve in Figure 5 (b) represents the data loss curve during the trading process. While the epoch number increases trading and validation curve decreases. It demonstrates that the proposed model is minimizing errors while it learns.



a) b)

Fig. 5. (a) accuracy graph and (b) loss graph of the proposed model visualization

3.2. Proposed CNN model evaluation

In this section, we evaluate the performance of our proposed model using various metrics such as precision, recall, and F-1 score. Those are used to assess our proposed model's classification capability of our dataset each class. A confusion matrix is also used to represent the distribution of real and predicted class labels visually.

The evolution metrics are:

- Precision
- Recall
- F1-Score
- Confusion Matrix

Precision: It indicates how many instances predicted in a particular class belong in that class. In other words, it measures the accuracy of true positives.

Recall: Measures the proportion of positive instances that the model correctly predicted.

F1-Score: This is the mean of precision and recall. It balances precision and recall, and it is helpful for an imbalanced class.

Table 4. Performance Evaluation of our proposed model

Class	Precision	Recall	F1 - Score
seborrheic keratosis	0.33	0.19	0.24
actinic keratosis	0.50	0.56	0.53
vascular lesion	0.40	0.12	0.19
pigmented benign keratosis	0.29	0.25	0.27
melanoma	0.36	0.88	0.51
basal cell carcinoma	0.50	0.56	0.53
dermatofibroma	0.00	0.00	0.00
nevus	0.31	0.25	0.28
squamous cell carcinoma	1.00	0.67	0.80

In Table 4, the precision values indicate how accurate the optimistic predictions are for each class. For example, for *squamous cell carcinoma*, a precision of 1.00 means that all predictions made for this class were correct, with no false positives. On the other hand, recall values show the proportion of actual instances that the model successfully identified. In the case of *squamous cell carcinoma*, a recall of 0.67 indicates that the model identified 67% of this class's cases. F1- scores provide a balanced measure by considering both precision and recall. A higher F1-score, such as 0.80 for *squamous cell carcinoma*, suggests a better trade-off between making accurate optimistic predictions and capturing a significant portion of the actual instances.

Table 4 demonstrates that multiple classes still need enhanced results in their classification segment. The model succeeded in recognizing squamous cell carcinoma and melanoma with promising accuracy yet exhibited low performance in classifying dermatofibroma and vascular lesions. The model demonstrates difficulties when effectively applying its learning to all categories. Additional training iterations, hyperparameter tuning, data augmentation, or exploration of alternative model architectures are encouraged to improve the overall performance.

Table 5. The Confusion Matrix proposed model

Class	seborrheic keratosis	actinic keratosis	vascular lesion	pigmented benign keratosis	melanoma	basal cell carcinoma	dermatofibroma	nevus	squamous cell carcinoma
seborrheic keratosis	2	0	1	0	9	3	0	1	0
actinic keratosis	1	6	0	1	0	5	0	2	1
vascular lesion	1	2	4	3	2	2	0	2	0
pigmented benign keratosis	0	0	0	2	9	5	0	0	0
melanoma	0	0	0	1	12	3	0	0	0
basal cell carcinoma	0	2	1	0	1	9	0	3	0
dermatofibroma	0	1	0	2	0	0	0	0	0
nevus	0	2	3	1	3	3	0	3	1
squamous cell carcinoma	0	1	0	0	0	0	0	0	2

Table 5 represents the Confusion matrix of our proposed model. It shows its performance. Here, each cell represents the number of predictions made by the model. It indicates if the model accurately classified the classes. The rows and columns correspond to the cancer type. Because this paper works with nine kinds of skin cancer, the confusion matrix is 9 × 9. The overall validation accuracy of 87% reported in this study was computed automatically during model training and evaluation on the entire validation dataset at each epoch. This metric represents the ratio of correctly predicted samples to the total validation samples, providing a global measure of model performance. In contrast, the confusion matrix presented was generated separately using the confusion matrix function. This matrix offers a detailed breakdown of the model's predictions by class and may be based on a subset of the validation data or outputs from an earlier training checkpoint. Consequently, the accuracy derived directly from this confusion matrix may differ from the overall reported accuracy.

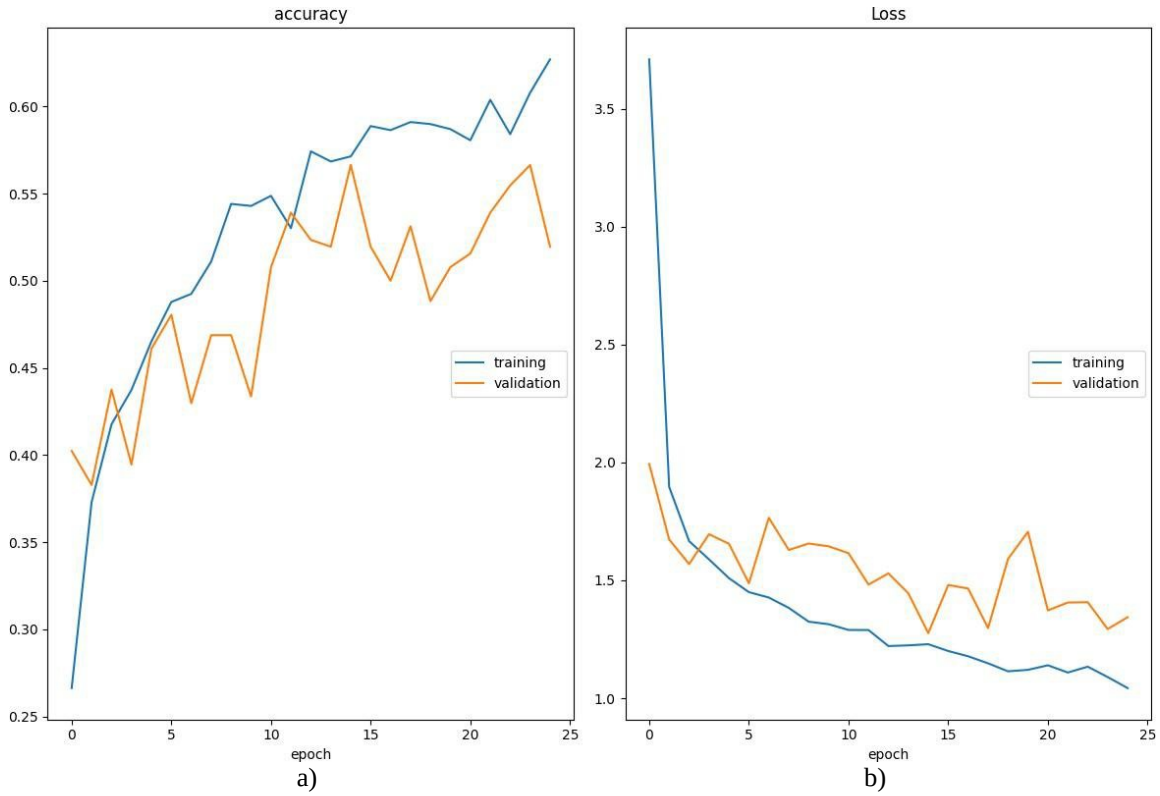
3.3. VGG16 Result

Various analyses evaluated the performance of the VGG16 model. In this section, we explain the key findings observed during the model's training and validation phases.

Figure 6 shows the visualization and changes in accuracy and loss during the training phase. The x-axis represents the number of training epochs, and the y-axis represents the corresponding accuracy or loss value.

The accuracy curve in Figure 6 (a) shows the model's ability to classify skin cancer images correctly. This graph shows that as the number of epochs increases, the training and validation accuracy also increase. This graph shows that this model has no overfitting because the training and validation curves are close. Figure 6 (a) shows that this model can classify with 65% accuracy.

The loss curve in Figure 6 (b) represents the data loss curve during trading. As the epoch number increases, the trading and validation curves decrease. This demonstrates that the proposed model minimizes errors while it learns.



a) b)

Fig. 6. (a) accuracy graph and (b) loss graph of VGG16 model visualization

In this section, we present and compare the outcomes of our proposed two models: the CNN and the VGG16 architecture. Our main objective was to analyze their capabilities in accurately classifying skin cancer.

The proposed CNN model showcased outstanding results, achieving an impressive accuracy rate of 87% on the validation set after 20 training epochs. Figure 5 (a) visually depicts this performance, showing the model's ability to classify skin cancer images accurately. Additionally, the curve's trend indicates effective generalization without overfitting. This promising outcome also shows the loss curve in Figure 5 (b), which shows that our model successfully minimized errors throughout its training process.

On the other hand, the VGG16 model shows a lower accuracy of 65% on the validation set after undergoing the same 25 epochs. The accuracy curve depicted in Figure 6 (a) shows a similar pattern in training and validation data, indicating the model's ability to generalize effectively to new samples.

Moreover, the loss curve in Figure 6 (b) signifies efficient error reduction during training, but this model has some overfitting.

To further evaluate the VGG16 model beyond accuracy and loss curves, we also analyzed its classification performance using standard metrics. The overall F1-score from the model indicated an average performance because it successfully maintained precision and recall levels across all categories. The confusion matrix showed that VGG16 performed well with selected classes but showed unsuccessful results when identifying other courses, leading to numerous classification errors. The evaluation results deliver an expanded knowledge of VGG16 model properties, thus making it easy to compare with the developed CNN model.

4. Discussion

The effective discovery of skin cancer and precise medical categorization contribute significantly to better patient survival and treatment success. The current diagnostic methods of expert analysis on biopsy specimens, dermatological examinations, and dermoscopic reviews demand much time and specialized interpretation capacity. Artificial intelligence (AI) with deep learning technologies shows substantial diagnostic potential in medical image analysis through its ability to improve accuracy and speed up medical diagnosis. We analyzed deep learning models with a customized CNN and VGG16 architecture to classify skin cancer types. The experimental findings showed that deep learning solutions generate superior outcomes than traditional machine learning methods through their strong accuracy and precision metrics, recall metrics, and F1-score values. The CNN-based model demonstrated better validation set performance, becoming the most efficient method for identifying skin cancer lesions. The CNN model performed better than VGG16 because it achieved 87% accuracy, while the other only managed 65%. The skin cancer classification tasks benefit from the CNN model because it extracts hierarchical features from images to generate better prediction results. The visual characteristics of different skin lesion types led both detection models to make wrong predictions despite their ability to classify such lesions correctly. The model encountered difficulties generalizing all skin lesion types because some categories contained insufficient training examples. The dataset requires a better balance between categories to enhance deep learning model optimization. Also, less accurate than the CNN model, the VGG16 model maintained dependable results throughout various classification scenarios. The VGG16 model demonstrates reliability when classifying different types of lesions, which shows its worth in particular cases requiring interpretability and class consistency. Our analysis indicates that CNN provides superior outcomes in detecting skin cancer because it performs better in accuracy and managing intricate features. The essential role of image preprocessing involved three steps: resizing, normalization procedures, data augmentation methods that included flips and rotations, and contrast adjustments to boost model effectiveness. The techniques implemented enhanced data consistency and increased diversity while concurrently reducing overfitting effects, making the models operate optimally on previously unseen pictures. Healthcare image analysis faces ongoing obstacles because of unbalanced classes, inconsistent image qualities, and inconsistent illumination and signal noise. Implementing advanced preprocessing methods, including histogram equalization and adaptive contrast adjustment, can address the existing issues while enhancing model performance outcomes. The unclear functioning of deep learning models is an immense barrier to their acceptance within clinical practice. Deep learning models, especially CNN, have displayed solid accuracy performance, yet their systematic decision-making algorithms remain unclear to interpret. Healthcare demands complete transparency regarding diagnostic decision-making because clinical staff must validate and trust model-generated diagnoses. Future developments should concentrate on creating explainable AI methods that demonstrate internal model operations so healthcare professionals will perceive them as acceptable. AI models need datasets that include diverse skin tones along with ethnic backgrounds and multiple skin lesion variations to guarantee their expanded applicability and fairness. Existing record datasets contain limited demographic representation, which generates possible prejudice in predictive modeling results. Using diverse datasets in deep learning model training helps achieve improved accuracy that leads to equal healthcare results. Our research proves CNNs are effective for initial skin cancer diagnosis, yet additional studies are needed to improve their deployment in medical practices. These include managing data variability, enhancing model explain ability, and optimizing computational efficiency. The successful clinical application of deep learning CNN models needs a solution to technical platform integration issues and practical deployment barriers. Success depends on further advancements in AI diagnostic technology. It holds exceptional promise to transform skin cancer detection through quicker results, better accuracy, and broader access to early identification for worldwide patients.

5. Conclusion

This study aimed to investigate how Convolutional Neural Networks (CNNs) operate with VGG16 architectural elements to detect skin cancer. The proposed CNN model obtained better classification outcomes since it achieved 87% accuracy, which surpassed the VGG16 model's accuracy of 65%. CNN is more suitable than other methods because it effectively detects complex skin cancer traits, thus becoming the principal selection for this scientific study. The VGG16 model displayed solid abilities in adapting to varied class distributions, yet its poor performance proves that the proposed CNN model should be preferred for this task. It has been found that deep learning improves the accuracy and efficiency of spotting skin cancer. According to our findings, a custom CNN can be more accurate than VGG16 when dealing with complex tasks in medical classifications. Since the CNN model achieved high accuracy, it shows that it can be practical for use in clinics as well as in remote settings. Emphasizing data preprocessing and equalizing classes enabled our models to do better in different situations and had a positive effect on their generalization. Being able to identify nine unique types of skin cancer shows that the model is useful in medical practice and supports accurate diagnosis. Ultimately, this work helps create AI-based solutions that doctors and healthcare professionals can use to detect cancer earlier, regardless of the background of patients. The implementation success of CNN alongside VGG16 reinforces deep learning's potential to generate rapid, accurate diagnostic aids for dermatologists, which promotes early detection and enhanced treatment results. Real-life deployment requires addressing three crucial problems, including unbalanced classes

in datasets, while improving interpretability features and expanding applicability variations among different populations. For future investigations, researchers should work on creating better balanced data collections and integrating explainable artificial intelligence (XAI) methods while adopting newer architectures like EfficientNet and Vision Transformers (ViTs). Modern skin cancer detection technology advancements will create better accessible and reliable early detection tools to revolutionize clinical practice.

Authorship Contributions

Taskin Sabit led the study, drafted the original manuscript, and conducted data visualization. His initiative and early contributions laid the foundation for the research work.

Faiza Tasnim contributed to model training and writing the results. She also enhanced the presentation of the paper through careful review.

Sadia Afrin Sara skillfully wrote the discussion and conclusion sections, providing critical insights that strengthened the overall narrative of the paper.

Sharia Tasnim Adrita composed the introduction, and **Maisha Tarannum** contributed significantly by meticulously formatting the text, verifying all references for accuracy.

Taskin Sabit and Faiza Tasnim collaborated closely to finalize the manuscript, ensuring consistency, coherence, and academic rigor throughout the paper.

All authors were actively involved in reviewing and approving the final version of the manuscript. Each member's collaboration and shared dedication greatly contributed to the success of this study

Declaration of conflicting interests

The authors declare no competing interests.

Funding

No financial support. This is an independent study and is not affiliated with any of the institutions mentioned.

Ethics

There are no ethical issues with the publication of this manuscript.

References

1. Haque, T., Rahman, K. M., Thurston, D. E., Hadgraft, J., & Lane, M. E. (2015). Topical therapies for skin cancer and actinic keratosis. *European Journal of Pharmaceutical Sciences*.
2. Naves, L. B., Dhand, C., Venugopal, J. R., Rajamani, L., Ramakrishna, S., & Almeida, L. (2017). Nanotechnology for the treatment of melanoma skin cancer. *Progress in Biomaterials*.
3. Karadag, A. S., & Parish, L. C. (2018). The status of the seborrheic keratosis. *Clinics in Dermatology*, *36*(2).
4. Bath-Hextall, F., Bong, J., Perkins, W., & Williams, H. (2004). Interventions for basal cell carcinoma of the skin: Systematic review. *British Medical Journal*.
5. Dika, E., Scarfì, F., Ferracin, M., Broseghini, E., Marcelli, E., Bortolani, B., … Lambertini, M. (2020). Basal cell carcinoma: A comprehensive review. *International Journal of Molecular Sciences*.
6. Clayman, G. L., Lee, J. J., Holsinger, F. C., Zhou, X., Duvic, M., El-Naggar, A. K., … Lippman, S. M. (2005). Mortality risk from squamous cell skin cancer. *Journal of Clinical Oncology*, *23*(4).
7. Alves, J. V. P., Matos, D. M., Barreiros, H. F., & Bártolo, E. A. F. L. F. (2014). Variants of dermatofibroma - A histopathological study. *Anais Brasileiros de Dermatologia*, *89*(3).
8. Rezaoana, N., Hossain, M. S., & Andersson, K. (2020). Detection and Classification of Skin Cancer by Using a Parallel CNN Model. In *Proceedings of 2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering, WIECON-ECE 2020*.
9. M, V. M. (2019). Melanoma Skin Cancer Detection using Image Processing and Machine Learning. *International Journal of Trend in Scientific Research and Development*, *Volume-3*(Issue-4).
10. Zghal, N. S., & Derbel, N. (2018). Melanoma Skin Cancer Detection based on Image Processing. *Current Medical Imaging Formerly Current Medical Imaging Reviews*, *16*(1).
11. Taye, M. M. (2023). Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions. *Computation*.
12. LeCun, Y., Hinton, G., & Bengio, Y. (2015). Deep learning (2015), Y. LeCun, Y. Bengio and G. Hinton. *Nature*, *521*.
13. ISIC Challenge ISIC-Archive.com . (n.d.). Retrieved March 26, 2025, from https://challenge**.**isic- archive.com/data/.