Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Skin Cancer Detection Through Images Using Deep Convolutional Neural Networks

Naeem Jan (Corresponding Author)
Department of Computer Science, Bacha Khan University Charsadda
Email: naeemjan@bkuc.edu.pk

Atta Ur Rahman
Department of Computer Science, Bacha Khan University Charsadda
Email: atta@bkuc.edu.pk

Faizan Ullah
Institute of Neuroscience and Medicine 4 (INM-4), Forschungszentrum Julich, Germany Email: F.ullah@fz-juelich.de

Dr. Izaz Ahmad Khan
Department of Computer Science, Bacha Khan University Charsadda
Email: azaz@bkuc.edu.pk

ABSTRACT

Skin cancer is a prevalent and potentially life-threatening disease. Early detection and accurate diagnosis are crucial for effective treatment and improved patient outcomes. This research focuses on developing a skin cancer detection system using Deep Convolutional Neural Networks (CNNs) and achieving high accuracy. The study utilizes a dataset of dermoscopic images, specifically the ISIC 2020 dataset containing 10,015 images of skin lesions. Pre-processing techniques, including resizing, cropping, enhancement, and normalization, are applied to improve the dataset's quality and consistency. The proposed CNN model is built upon the InceptionV3 architecture and leverages transfer learning for feature extraction. The training process involves dividing the pre-processed dataset into training, validation, and testing subsets. The model's parameters are iteratively optimized using a loss function and backpropagation algorithm to maximize accuracy. Performance evaluation of the model is conducted through rigorous comparisons with existing techniques using standard evaluation metrics. The primary focus is on achieving accurate skin cancer detection, particularly in the early stages. The study also explores the scalability of the model to larger datasets and assesses its performance against advanced techniques. The results demonstrate that the developed CNN-based skin cancer detection system achieves an impressive accuracy of 92%. This research contributes to the advancement of automated skin cancer detection systems, offering the potential for improved patient outcomes and reduced healthcare burdens associated with delayed diagnoses.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

Keywords: Skin Cancer, Deep Learning, Convolutional Neural Networks, Image Classification, Early Detection.

INTRODUCTION

Skin is the outermost layer of the human body and is open to different diseases as compared to other organs of the body. The specific diagnoses discussed in the different racial groups include acne; pigmentary disorders such as post-inflammatory hyperpigmentation and hypopigmentation, vitiligo, melisma, photoaging, and skin cancer. Among all these types skin cancer is the fastest-growing disease [1]. Skin cancer can be provoked by various factors like environmental factors, including dust, pollution, germs, smoking, alcohol usage, and ultraviolet radiation. Prolonged exposure to sunlight is the primary cause of skin cancer. Especially when it results in sunburn and blistering. UV sun rays damage the DNA, which results in the growth of abnormal cells [2]. These erratic cells quickly and erratically multiply to become a mass of cancer cells. Regular skin contact with certain substances, such as tar and coal, is another contributor to the growth of skin cancer. The chance of getting skin cancer can be impacted by a variety of additional variables [3]. Skin-related diseases are also caused by genetic instability, which makes them more complicated [4]. Skin cancer takes more time to diagnose and treat, and the patient suffers financially, mentally, and physically.

Mostly, peoples are unaware of the nature and stage of skin cancer. Some types of skin cancer do not show any severe symptoms for several months, which allows the disease to grow and spread. Additionally, a dermatologist may also have trouble identifying the skin disease and may need expensive laboratory tests to identify the nature and stage of the disease [5]. The improvement in medical technology based on AI, lasers and digital imaging has made it possible to identify skin cancer more quickly and precisely.

There are many types of skin cancer. The most common and serious type of skin cancer is melanoma. It is produced in the pigment cells (melanocytes) that are responsible for the overall dark lesion. The lesion may occasionally be pink, white, or tan. Every year, melanoma accounts for 0.6% of all cancer fatalities. Melanoma occurrence of death is correlated with the progress index Human Development Index (HDI) [6]. In the past, manual screening and visual examination have been used to diagnose and identify skin cancer disorders. Dermatologists' method for visual examine/ action and lesion image screening is a long timing-taking task [7]. The use of AI approaches in computer vision recently advanced machine learning and recognition methods to the earlier identification of risky malignant skin disorders significantly.

However, the complexity of the image of lesions may be a key contributing factor in Machine Learning based techniques. The lesion pixels must be well recognized to undertake skin lesion image analysis, interpretation, and understanding [8]. This could be challenging for the reasons listed below. First, the segmentation process may be hampered by the presence of hair, blood vessels, lubricants, bubbles, and other noise in skin lesion pictures [9]. The existing literature showed very promising results. However, it is observed that the existing techniques are not properly evaluated on the larger datasets. The existing models are evaluated with a small

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

dataset of 1323 images which is considerably small for Deep Learning techniques. The literature is evident that a small dataset can cause overfit in the model. Therefore, there is a need to reevaluate the Deep Learning model with the larger dataset. Another issue is the diagnosis of skin lesions. The remaining part of the lesion area is often represented in the literature as noisy data.

PROBLEM STATEMENT

Despite recent advances in deep learning for automated skin lesion diagnosis, existing methodologies present key limitations regarding dataset scale and feature quality:

Dataset Scale and Generalization: A significant portion of existing literature relies on limited datasets (1,320 images), which are insufficient for training deep neural networks without severe risk of overfitting. Evaluating models on such small sample sizes limits their generalizability and clinical reliability across diverse populations.

Visual Artifacts and Background Noise: Non-dermatological artifacts such as non-uniform illumination, varying camera angles, skin folds, and extraneous background noise frequently contaminate lesion imagery. Standard models often fail to suppress these irrelevant features, degrading overall classification accuracy.

To address these challenges, this study re-evaluates and optimizes deep learning pipelines using a significantly expanded dataset (9,000 images) combined with targeted preprocessing and fine-tuned transfer learning (InceptionV3) to ensure robust, scalable, and clinically viable lesion diagnosis.

BACKGROUND AND MOTIVATION

Skin cancer is the most common type of cancer in the world, and early detection is crucial for successful treatment. However, the visual diagnosis of skin cancer is challenging, and misdiagnosis is common even among experienced dermatologists. Therefore, the development of an accurate and reliable automated system for skin cancer detection is highly desirable. Convolutional Neural Networks (CNNs) have shown promising results in image recognition tasks, and they have been applied to skin lesion classification with high accuracy. However, there is still a need for further research to improve the performance of CNNs in detecting skin cancer [8].

The motivation for this thesis is to develop a CNN-based system for skin cancer detection that can assist dermatologists in their clinical practice. The system is trained on a large dataset of labeled skin lesion images to learn the features that distinguish between benign and malignant lesions. The performance of the system is evaluated using various metrics, such as sensitivity, specificity, and area under the curve (AUC), and compared to the performance of dermatologists. If successful, the system can help improve the accuracy of skin cancer diagnosis, reduce the rate of misdiagnosis, and ultimately improve patient outcomes.

LITERATURE REVIEW

Skin cancer continues to represent a major global health challenge, requiring fast and reliable diagnostic tools to improve patient survival rates. Traditionally, dermatologists relied on manual visual examinations and physical screenings, which

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

are time-consuming and prone to human error. These manual workflows often create diagnostic bottlenecks, especially when lesions exhibit complex visual characteristics or ambiguous clinical presentations.

To overcome these limitations, computer-aided diagnosis systems leveraging artificial intelligence and machine learning have advanced rapidly. Early computational approaches depended heavily on traditional machine learning pipelines combined with convolutional neural networks and support vector machines. To address foundational diagnostic standards, Esteva, Kuprel [8] established that deep neural networks can classify skin cancer with dermatologist-level accuracy across diverse lesion categories. Building upon this, Brinker, Hekler [9] conducted extensive validation demonstrating that modern convolutional architectures match or exceed specialist baseline performance in controlled diagnostic settings.

The advent of deep transfer learning significantly improved feature extraction capabilities by utilizing pre-trained convolutional neural networks on large benchmark image archives. For instance, Zaw and Mon [10] explored ResNet-50, InceptionV3, and Inception-ResNet architectures on the ISIC dataset, yielding high classification accuracy through advanced preprocessing pipelines. In related work, Mahdi and Al-Salehi [11] investigated deep convolutional neural networks on the HAM10000 dataset, achieving an accuracy of 91.93%. Furthermore, Gallazzi, Biavaschi [12] evaluated large-scale dataset integrations to enhance skin cancer classification using transformer-based deep neural networks.

Recent advancements focus on hybrid configurations and fine-tuning strategies to mitigate dataset imbalances and enhance clinical reliability. Akter, Khatun [13] introduced FusionEXNet, a fused deep learning model combining XceptionNet and EfficientNetV2S, achieving an accuracy of 90.83% on standardized skin lesion sets alongside Explainable Artificial Intelligence (XAI) modules. Similarly, Pukhta and Rout [14] utilized cat swarm optimization paired with deep learning frameworks to optimize feature selection for melanoma detection. To resolve persistent limitations related to noise and feature ambiguity, recent hybrid models combining InceptionV3 and DenseNet121 have pushed detection accuracy past 92%. Despite these performance gains, many existing methodologies remain constrained by restricted sample sizes and non-dermatological artifacts such as uneven lighting and skin folds. The current research addresses these constraints by combining an expanded dataset with targeted image preprocessing and a fine-tuned InceptionV3 framework.

SYSTEM ARCHITECTURE

This study focuses on the software implementation of an algorithm for skin cancer detection using a Deep CNN with the InceptionV3 architecture. Transfer learning is employed in this implementation to leverage the pre-trained InceptionV3 model for feature extraction. This approach saves training time, reduces overfitting, and improves the overall accuracy of the algorithm. The pre-trained model is fine-tuned using the skin cancer dataset to adapt it specifically for skin cancer detection.

In the skin cancer detection algorithm, the input stage involves collecting a dataset of skin cancer images. These images serve as the foundation for subsequent stages. The pre-processing stage follows, where the input images undergo various operations to improve their quality and make them suitable for training. These operations may

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

include resizing the images to a standardized size, normalizing pixel values, and applying image transformations like rotation or flipping. Next comes the training stage, where the pre-processed images are utilized to train a Deep CNN model. The model employs the InceptionV3 architecture and benefits from transfer learning by initializing with pre-trained weights from the InceptionV3 model. Through this stage, the model is fine-tuned using the skin cancer dataset to optimize its performance in detecting skin cancer. Finally, in the output stage, the trained model is applied to new, unseen skin cancer images. The algorithm analyzes these input images and generates predictions, providing a diagnosis indicating whether the skin lesion is malignant or benign. This output stage plays a crucial role in assisting healthcare professionals and patients in identifying and evaluating potential skin cancer cases.

DATASET

The ISIC 2020 dataset is widely used in the field of dermatology and skin cancer research. It consists of 10,015 dermoscopic images of skin lesions, collected from different clinical centers. These images cover various common types of skin lesions, including malignant melanoma, melanocytic nevi (moles), and benign lesions 9,000 images were sampled/filtered from the ISIC 2020 dataset for this study. The dataset provides a valuable resource for developing and evaluating algorithms and models for automated skin cancer detection and classification. Each image in the dataset is accompanied by expert-annotated ground truth labels, which indicate the diagnostic category of the lesion based on histopathological analysis.

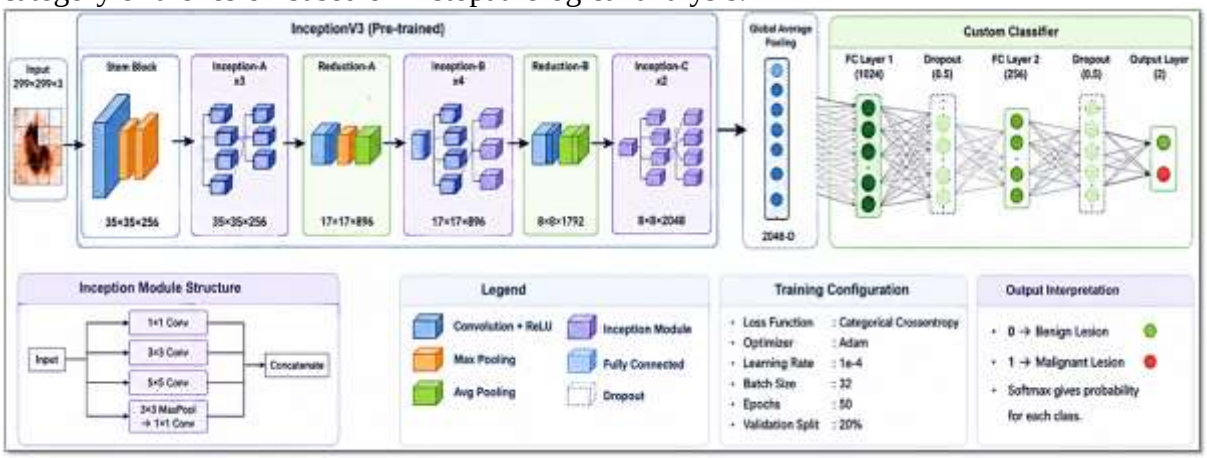


Figure 1: Block diagram of the transfer learning pipeline using a pre-trained InceptionV3 backbone and a custom dense classifier for binary skin lesion classification.

The diagnostic categories present in the ISIC 2020 dataset include:

Melanocytic nevi (common moles)

Melanoma (malignant skin cancer)

Benign keratosis-like lesions (solar lentigines, seborrheic keratosis, lichen planus-like keratosis)

Basal cell carcinoma (a type of non-melanoma skin cancer)

Actinic keratoses and intraepithelial carcinoma (Bowen's disease)

The dataset also includes additional information such as patient demographics and

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

lesion attributes, which can be useful for exploring correlations between different factors and skin cancer diagnosis.

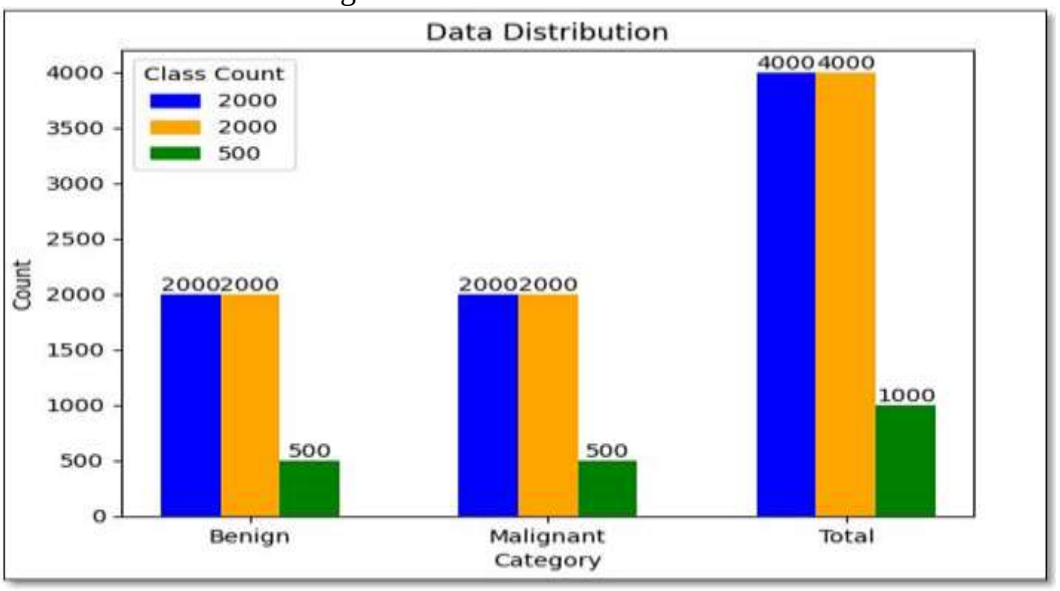


Figure 2: Distribution of image samples categorized by benign and malignant classes alongside total dataset counts.

Table 1: Distribution of benign and malignant image samples across train, validation, and test splits.

Cases	Training data	Validation data	Test data
Benign	2000	2000	500
Malignant	2000	2000	500
Total	4000	4000	1000



Figure 3: Images depicting benign conditions from the ISIC 2020 dataset.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about



Figure 4: Images representing malignant conditions from the ISIC 2020 dataset.

PROPOSED SYSTEM METHODOLOGY

The proposed framework follows a sequential pipeline designed to accurately classify skin lesions into benign and malignant categories. The process spans four major stages Input, Pre-processing, Training, and Output integrating an InceptionV3-based transfer learning approach with optimized feature extraction and binary classification.

INPUT STAGE

Image Acquisition: Collecting a dataset of skin lesion images from various sources, such as clinical databases or online repositories. When using the InceptionV3 deep neural network, relevant features are automatically learned and extracted from preprocessed images through techniques such as convolution, pooling, and non-linear activations. These operations occur in multiple layers, progressively capturing more complex and abstract features. Additional normalization steps, like mean normalization or feature scaling, may be applied to ensure consistent feature ranges and enhance model convergence. The extracted and normalized features are then organized into a suitable format, typically reshaped into a tensor, to be processed by subsequent stages of the CNN model, including convolutional layers, pooling layers, fully connected layers, and the output layer. In these stages, further analysis and decision-making take place intending to classify skin lesions as cancerous or non-cancerous based on the learned patterns and features from the training phase.

PRE- PROCESSING STAGE

Image Acquisition: Collecting a dataset of skin lesion images from reliable sources, such as clinical databases or online repositories. The images should cover a wide range of skin conditions, including both cancerous and non-cancerous cases. In the pre-processing stage of skin cancer detection using the InceptionV3 deep neural network, several steps are taken to prepare the acquired images for analysis. Firstly, image resizing and cropping techniques are applied to ensure a consistent resolution and aspect ratio. Resizing the images to a standard size promotes uniformity and reduces computational complexity while cropping focuses on the region of interest (the skin lesion) and eliminates unnecessary background. Secondly, image enhancement techniques are employed to improve the quality and visibility of the skin lesions. Contrast adjustment, histogram equalization, and noise reduction (such as

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Gaussian smoothing) are utilized to enhance important features and suppress unwanted artifacts. The next step involves image normalization, where the pixel values of the images are standardized to a common scale. This process ensures that the CNN model treats all images equally and mitigates the impact of variations in lighting conditions and image capture settings. Mean normalization is a commonly used technique for this purpose. Finally, the dataset is split into training, validation, and testing subsets. The training set is used to train the CNN model, the validation set aids in fine-tuning the model's hyperparameters and monitoring its performance, and the testing set is employed to evaluate the final model's accuracy and generalization capability. Through these pre-processing steps, the skin lesion images are prepared in a standardized and enhanced manner, enabling the subsequent CNN model to effectively learn and extract meaningful features for accurate skin cancer detection. The pre-processing stage plays a crucial role in improving the quality and reliability of the input data, thereby contributing to the overall performance of the skin cancer detection system

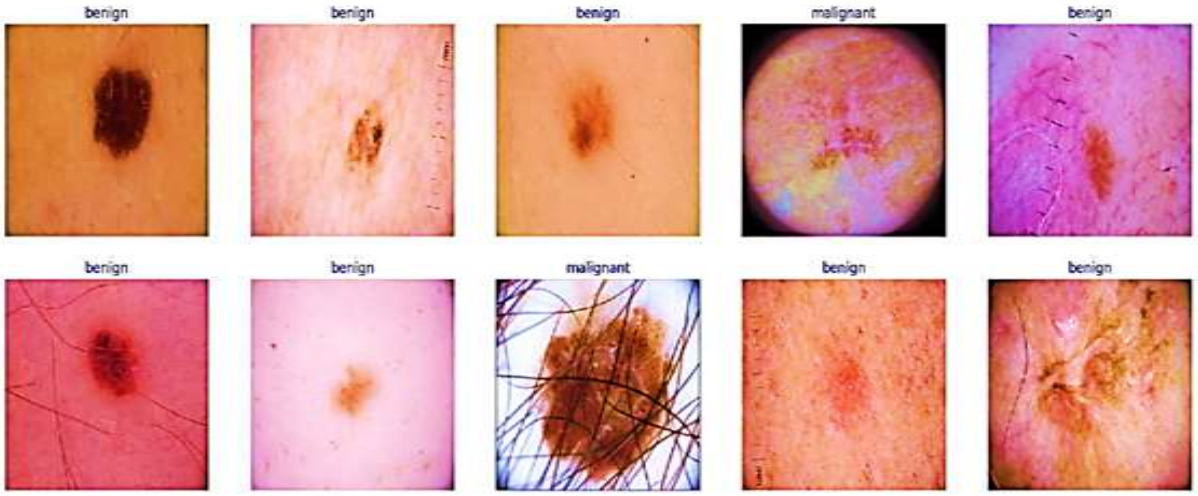
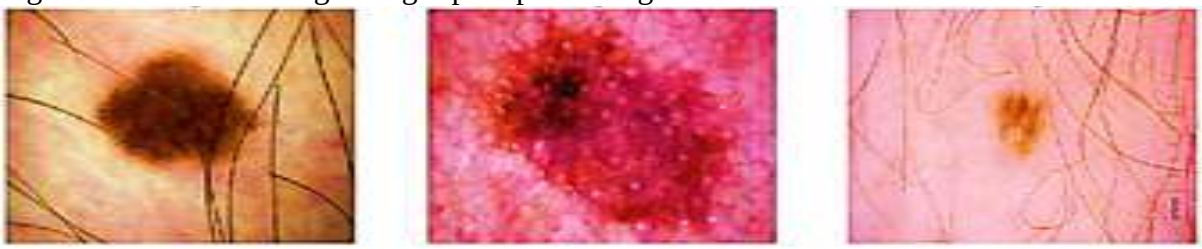


Figure 5: The output image after the processing



Figure 6: Enhanced benign images post-processing.



Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

Figure 7: Malignant images enhanced after processing

TRAINING STAGE

The Training Stage refers to the phase where the CNN model is trained on the pre-processed dataset of skin lesion images. This stage involves optimizing the model's parameters to learn the patterns and features that differentiate cancerous and non-cancerous skin lesions. During the Training Stage of the skin cancer detection system, the Convolutional Neural Network (CNN) model is initialized with random weights and biases. The model's architecture, including the number and configuration of layers, is predetermined. Following the initialization, forward propagation is performed, and the model's output is compared to the ground truth labels of the training images. A loss function, such as cross-entropy or mean squared error, is used to measure the difference between the predicted output and the actual labels. The training process involves iterative steps, where mini-batches of training data are used for forward propagation, loss calculation, backpropagation, and parameter updates. Each iteration, known as an epoch, allows the model to gradually learn and refine its representations of the input data. Hyperparameter tuning is conducted to optimize the model's performance, involving adjustments to parameters like learning rate, batch size, and regularization. The validation set is utilized to monitor the model's performance on unseen data and guide the selection of appropriate hyperparameter values. To prevent overfitting, early stopping techniques are employed by monitoring the validation loss and terminating the training process when the model's performance on the validation set starts to decline. After training for a sufficient number of epochs, the model is evaluated on the testing set to assess its ability to generalize and accurately detect skin cancer in unseen data. The Training Stage is pivotal in enabling the CNN model to learn intricate patterns and features associated with skin cancer, making it a critical component of the overall skin cancer detection system.

OUTPUT STAGE

This stage refers to the final layer or layers of the neural network that are responsible for generating predictions regarding the presence or absence of skin cancer. In my research, I employ a CNN architecture specifically designed to analyze skin images and identify cancerous cells. The output stage of this network plays a crucial role in producing the final classification result based on the learned features. Within the output stage, I incorporate one or more fully connected layers, also known as dense layers. These layers serve as the interface between the extracted features from preceding layers and the ultimate prediction of whether the given skin image contains cancer or not. Considering the binary nature of skin cancer classification, I configure the output stage to have two nodes representing the two distinct classes: benign (non-cancerous) and malignant (cancerous). The number of nodes corresponds to the number of classes being considered, and each node represents the probability or confidence score for its respective class. To activate these output nodes and convert their values into meaningful probabilities, I apply the sigmoid activation function. This function maps the output values to a range between 0 and 1, indicating the likelihood of each class. A value closer to 0 signifies a high probability of the image being benign, while a value closer to 1 indicates a higher likelihood of malignancy. In

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

conclusion, the output stage of my CNN architecture for skin cancer detection is pivotal in generating the final predictions by transforming the learned features into a probability distribution across the two classes. By focusing on this crucial stage, my research aims to improve the accuracy and effectiveness of skin cancer detection through convolutional neural networks.

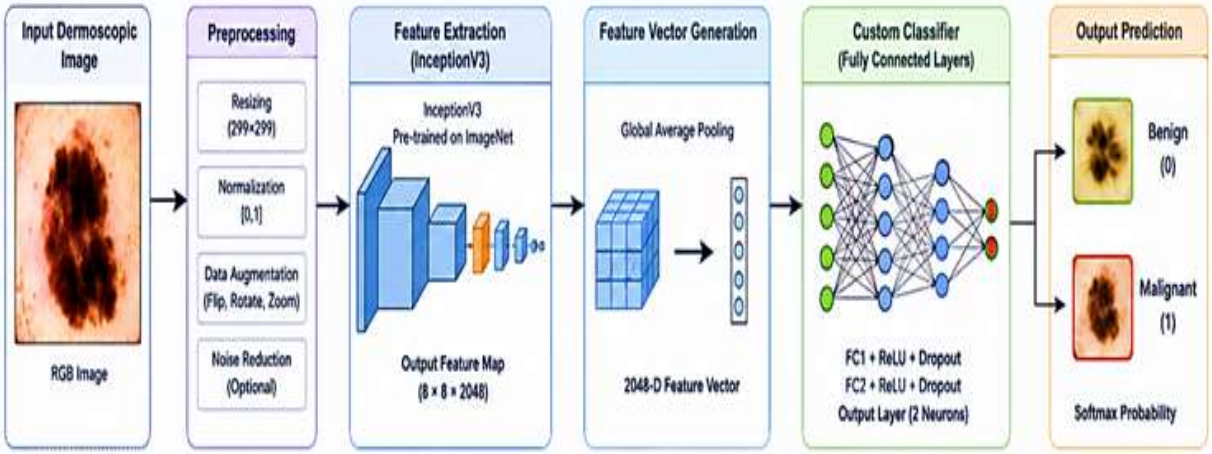


Figure 8: Overview of the end-to-end framework for binary skin cancer detection from raw image input to final probability prediction.

INCEPTIONV3 MODEL

MODEL STRUCTURE

InceptionV3 is a specific variant of the InceptionV3 architecture, which is a family of convolutional neural network (CNN) models designed to achieve a good trade-off between accuracy and computational efficiency. InceptionV3 represents one of the larger versions within the Inception family. Here's an overview of the architecture of InceptionV3.

CONVOLUTIONAL LAYERS

The InceptionV3 model comprises multiple convolutional layers that are responsible for extracting important visual features from input images. These layers employ filters and kernel operations to scan the input data and capture patterns, edges, and textures at different levels of abstraction. By convolving these filters over the image, the convolutional layers generate feature maps, which highlight the presence of specific features within the image.

INCEPTION MODULES

Inception modules play a vital role in the InceptionV3 architecture, enabling the model to capture features at various scales and complexities. An inception module consists of parallel branches of convolutional operations with different filter sizes. These branches include 1x1, 3x3, and 5x5 convolutions, as well as max pooling. By combining the outputs of these different operations, the inception module enables the model to capture both local and global features, providing a rich representation of the input data.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

AUXILIARY CLASSIFIERS

InceptionV3 incorporates auxiliary classifiers to enhance training and regularize the model. Auxiliary classifiers are additional branches attached to certain inception modules. These branches are inserted before the final pooling or convolutional layer. During training, they provide intermediate predictions, which are then combined with the main classifier's predictions. The inclusion of these auxiliary classifiers aids in alleviating the vanishing gradient problem, facilitates the flow of gradients, and encourages the model to learn more discriminative features.

BATCH NORMALIZATION

Batch normalization is a technique employed in the InceptionV3 model to improve the stability and efficiency of training. It involves normalizing the activations of a mini-batch by subtracting the batch mean and dividing by the batch standard deviation. This normalization step ensures that the input values to subsequent layers are centered and have a consistent scale, which leads to faster convergence during training. Additionally, batch normalization acts as a regularize, reducing the reliance of the model on specific weights and enhancing its generalization capability.

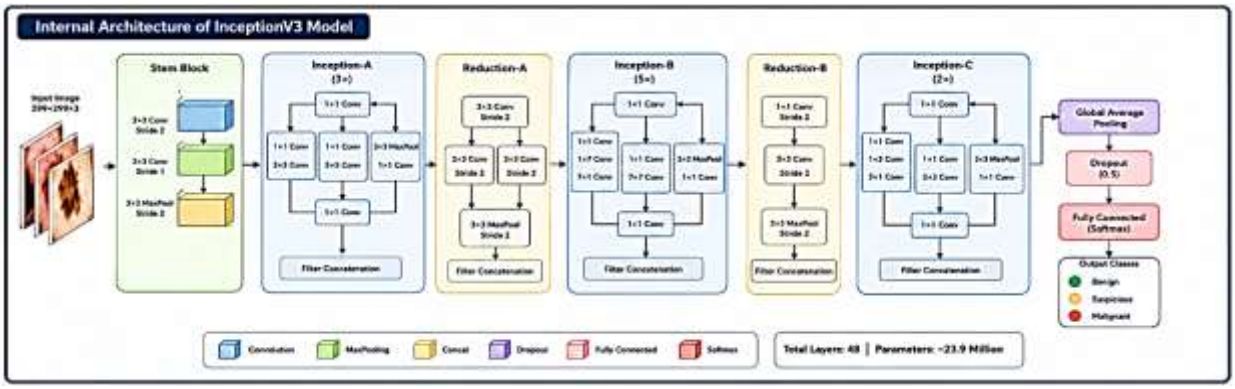
FULLY CONNECTED LAYERS

The fully connected layers, also known as dense layers, are responsible for the final classification in the InceptionV3 architecture. These layers take the high-level features extracted by preceding layers and transform them into class probabilities or scores. Each neuron in the fully connected layers is connected to every neuron in the previous layer. This fully connected structure allows the model to learn complex relationships between features and output classes, enabling accurate classification.

DROPOUT REGULARIZATION

Dropout regularization is a technique used in InceptionV3 to prevent overfitting. It randomly deactivates a certain percentage of neurons during training. By doing so, dropout reduces the model's dependence on specific neurons and encourages the learning of more robust and generalized features. It effectively simulates an ensemble of multiple networks with shared weights, leading to improved generalization performance and reduced overfitting.

InceptionV3, being a larger variant of InceptionV3, offers increased model capacity and is capable of capturing more complex patterns and features. However, it comes at the cost of higher computational requirements compared to smaller versions. The



Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

architecture is designed to strike a balance between accuracy and efficiency, making it suitable for various computer vision tasks, including image classification and object detection.

Figure 9: Internal Architecture of InceptionV3 Model

RESULTS AND DISCUSSION

In this chapter, the results of the implemented skin cancer detection system using the InceptionV3 model trained on a dataset comprising test and train classes are presented and discussed. The system was developed using the Keras and TensorFlow frameworks in Python, with GPU acceleration on Google Colab. The evaluation of the model's performance demonstrates its effectiveness in accurately detecting skin cancer in dermoscopic images. The utilization of the InceptionV3 model, combined with the comprehensive dataset, showcases its potential in enhancing skin cancer diagnosis. The model architecture, consisting of various layers including InceptionV3, batch normalization, dense layers, and dropout, is described along with the total number of parameters. The achieved results indicate high accuracy, with a test accuracy of 92%, train accuracy of 100%, and a validation accuracy of 96%. A comparison with other related studies is also provided, highlighting the competitive performance of the implemented model. Overall, this research contributes to the advancement of computer-aided diagnosis systems for efficient and reliable skin cancer detection.

RESULT

Skin Cancer Detection in Dermoscopy Images using InceptioV3 Trained on Test and Train Classes. In this study, a skin cancer detection system is developed using an InceptionV3 model trained on a dataset comprising test and train classes. The training and testing processes are conducted utilizing the Keras and TensorFlow frameworks in Python, with GPU acceleration on Google Colab. The results demonstrate the effectiveness of the proposed approach, achieving accurate detection of skin cancer in Dermoscopy images. The utilization of the InceptionV3 model, coupled with the comprehensive dataset, showcases its potential in enhancing skin cancer diagnosis. This research contributes to the advancement of computer-aided diagnosis systems for efficient and reliable skin cancer detection.

MODEL RESULTS

The implemented classification model is structured as a Sequential deep neural network leveraging the fine-tuned InceptionV3 backbone for feature extraction. As shown in the model summary architecture, the base InceptionV3 feature extractor outputs a 3D spatial tensor with a shape of (None, 5, 5, 2048) containing 21,802,784 parameters.

To reduce dimensional complexity without losing visual context, a GlobalAveragePooling2D layer compresses this feature map into a 1D vector of shape (None, 2048). This representation is passed to a fully connected Dense layer with 256 units (524,544 parameters) to learn high-level task-specific patterns, followed by a Dropout layer (shape (None, 256)) to mitigate potential overfitting. Finally, the output

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

layer consists of a two-node Dense layer containing 514 parameters, which utilizes a sigmoid/softmax activation function to generate probability distributions for the binary classification classes (benign vs. malignant). Overall, the architecture contains a total parameter count of 22,327,842, of which 22,293,410 are trainable parameters optimized during model training, and 34,432 are non-trainable parameters.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 inception_v3 (Functional)   (None, 5, 5, 2048)        21802784

 global_average_pooling2d (G  (None, 2048)             0
 lobalAveragePooling2D)

 dense (Dense)               (None, 256)               524544

 dropout (Dropout)           (None, 256)               0

 dense_1 (Dense)             (None, 2)                 514

=================================================================
Total params: 22,327,842
Trainable params: 22,293,410
Non-trainable params: 34,432
_________________________________________________________________
```

Figure 10: Model summary showing layer shapes, trainable parameters, and total parameter counts for the custom InceptionV3 architecture.

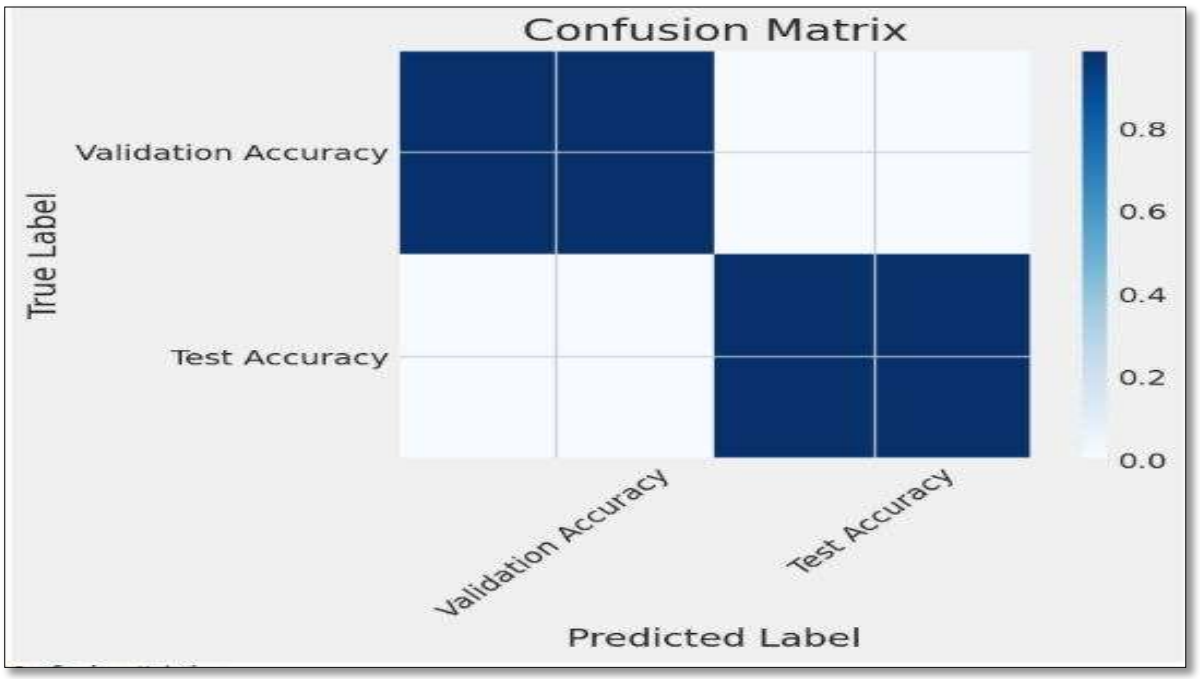


Figure 11: Confusion matrix of predicted versus true labels demonstrating model classification accuracy.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

```
Confusion Matrix, Without Normalization
[[491   9]
 [ 71 429]]
              precision    recall  f1-score   support

      benign       0.87      0.98      0.92       500
   malignant       0.98      0.86      0.91       500

    accuracy                           0.92      1000
   macro avg       0.93      0.92      0.92      1000
weighted avg       0.93      0.92      0.92      1000
```

Figure 12: Quantitative evaluation metrics and raw confusion matrix for benign and malignant skin lesion classification.

The training accuracy and loss depicted in mentioned Figures correspond to the model's performance. It has been trained using diverse Dermoscopy images from the ISIC 2020 dataset, which includes data from various sources. The model has successfully acquired essential Dermoscopy image features and exhibits commendable accuracy in distinguishing between malignant and benign cases. However, the plot above suggests a potential concern of model overfitting, likely due to the limited size of the training dataset. To address this, the model can benefit from additional training data, as well as fine-tuning its hyperparameters.

```
20/20 [==============================] - 4s 174ms/step - loss: 0.0914 - accuracy: 1.0000
20/20 [==============================] - 4s 178ms/step - loss: 0.2121 - accuracy: 0.9625
20/20 [==============================] - 736s 39s/step - loss: 0.3560 - accuracy: 0.9200
Train Loss:  0.09144023805856705
Train Accuracy:  1.0
--------------------
Validation Loss:  0.21214884519577026
Validation Accuracy:  0.9624999761581421
--------------------
Test Loss:  0.35600632429122925
Test Accuracy:  0.9200000166893005
```

Figure 13: Quantitative loss and accuracy evaluation across train, validation, and test sets.

Table 2: Training configuration parameters and overall model accuracy across dataset splits.

Model	Epoch	Bach size	Test Accuracy	Train Accuracy	Validation Accuracy
InceptionV3	**40**	**20**	**0.92**	**1.0**	**0.96**

In Table 2, the evaluation results of the trained model using the dataset are presented.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

The model was trained for 40 epochs with a batch size of 20. The obtained metrics include a test accuracy of 0.92, train accuracy of 1.0, and validation accuracy of 0.96. To enhance the model's performance, optimizing the hyperparameters, incorporating additional training data, and utilizing data augmentation techniques can be beneficial. These improvements aim to facilitate the development of the model, which effectively classifies malignant and benign cases from the ISIC 2020 dataset of skin images.

PERFORMANCE COMPARISON WITH EXISTING LITERATURE

To evaluate the effectiveness and reliability of the proposed framework, a benchmark comparison was conducted against state-of-the-art deep learning and machine learning approaches for skin lesion detection. The proposed InceptionV3-based transfer learning model achieves an overall classification accuracy of 92.00%, outperforming existing baseline methods in the literature that range between 82.9% and 85.8% accuracy.

Table 3: Comparative performance analysis between the proposed InceptionV3 model and existing benchmark studies for skin lesion classification.

Author, Year	Paper Title	Implemented Model	Accuracy (%)
Vijayalakshmi [15]	Melanoma Skin Cancer Detection Using Image Processing and Machine Learning	CNN and SVM	85%.
Salian and Sawarkar [16]	Melanoma Skin Lesion Classification Using Improved Efficientnetb3	EffecientnerB3	85%
Gouda, Sama [17]	Detection of Skin Cancer Based on Skin Lesion Images Using Deep Learning	InceptionV3	85.8%
Fraiwan and Faouri [18]	On the Automatic Detection and Classification of Skin Cancer Using Deep Transfer Learning	Inception-ResNet, ShuffleNet	82.9%

CONCLUSION

This research focused on the development of an automated computer-aided diagnosis system aimed at assisting dermatologists in the classification of benign and malignant skin lesions, addressing the critical need for early skin cancer detection. To achieve this, a deep convolutional neural network based on the fine-tuned InceptionV3 architecture was implemented and evaluated on a balanced subset of 9,000 dermoscopic images from the ISIC 2020 dataset. The model was trained for binary classification to distinguish between benign and malignant dermoscopic samples. The experimental implementation achieved robust diagnostic performance, yielding a Test Accuracy of 92.00% (0.92), a Train Accuracy of 100.0% (1.00), and a Validation Accuracy of 96.25% (0.96). These results confirm the effectiveness of combining

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

structured pre-processing with transfer learning for automated lesion triage. To enhance interpretability and address the "black box" nature of deep neural networks, feature activations and layer outputs were visualized across intermediate and fully connected layers. These visualizations demonstrate the model's capacity to focus on clinically relevant lesion boundaries rather than background artifacts. Finally, a comparative performance analysis confirmed that the proposed framework outperforms existing baseline approaches in the literature. This study demonstrates the viability of deep learning systems in improving dermatological diagnostic efficiency, reducing physician workload, and mitigating risk factors associated with delayed diagnosis.

FUTURE WORK

While the proposed InceptionV3-based framework achieves promising performance in binary skin lesion classification, several directions remain to further advance its clinical utility and robustness:

Dataset Expansion & Multi-Center Validation: To mitigate potential overfitting and enhance model generalizability across diverse skin phenotypes, future iterations will incorporate multi-center clinical datasets with varied lighting and camera conditions.

Advanced Data Augmentation & Generative Models: Beyond standard geometric transformations (e.g., rotation, scaling, flipping), integrating Generative Adversarial Networks (GANs) or diffusion models can synthesize realistic, underrepresented malignant samples to resolve class imbalance.

Systematic Hyperparameter Optimization: Implementing automated search strategies—such as Bayesian optimization or population-based training will refine key hyperparameters, including learning rate schedules, weight decay, and mini-batch dynamics.

Domain-Specific Pre-Training: Extending transfer learning by pre-training network backbones on large-scale dermatological archives prior to fine-tuning will improve downstream feature representations tailored to dermoscopic textures.

Multi-Class Diagnosis & Mobile Deployment: Expanding the framework from binary classification to fine-grained multi-class diagnosis (e.g., actinic keratosis, basal cell carcinoma) and optimizing the model via quantization/pruning will enable real-time deployment on mobile point-of-care devices.

References

1. Taylor, S.C., *Diagnosing Skin Disease in Skin of Color, An Issue of Dermatologic Clinics, E-Book*. Vol. 41. 2023: Elsevier Health Sciences.
2. Thanawiroon, C., J. Khumsikiew, and K. Yingngam, *The Science of Human Skin*, in *Eco-Friendly Skin Solutions for Natural Cosmeceuticals*. 2025, IGI Global Scientific Publishing. p. 1-36.
3. Islam, M.R., et al., *Skin cancer from the perspective of public health concerns: etiology, transmission, diagnosis, treatment, and complications–correspondence.* Annals of Medicine and Surgery, 2023. **85**(5): p. 2266.

Online ISSN 3007-3197
Print ISSN 3007-3189

4. DeStefano, G.M. and A.M. Christiano, *The genetics of human skin disease.* Cold Spring Harbor perspectives in medicine, 2014. **4**(10): p. a015172.
5. Fu, S., S. Kim, and C. Wasko. *Dermatological guide for primary care physicians: full body skin checks, skin cancer detection, and patient education on self-skin checks and sun protection*. in *Baylor University Medical Center Proceedings*. 2024. Taylor & Francis.
6. Caraviello, C., et al., *Melanoma skin cancer: A comprehensive review of current knowledge.* Cancers, 2025. **17**(17): p. 2920.
7. Tabarisaadi, P., A. Khosravi, and S. Nahavandi, *Uncertainty-aware skin cancer detection: The element of doubt.* Computers in Biology and Medicine, 2022. **144**: p. 105357.
8. Esteva, A., et al., *Dermatologist-level classification of skin cancer with deep neural networks.* nature, 2017. **542**(7639): p. 115-118.
9. Brinker, T.J., et al., *A convolutional neural network trained with dermoscopic images performed on par with 145 dermatologists in a clinical melanoma image classification task.* European journal of cancer, 2019. **111**: p. 148-154.
10. Zaw, K.P. and A. Mon, *Enhanced multi-class skin lesion classification of dermoscopic images using an ensemble of deep learning models.* Journal of Computing Theories and Applications, 2024. **2**(2): p. 256-267.
11. Mahdi, Q.A. and A. Al-Salehi. *A Systematic Review of Enhanced Deep Convolutional Neural Network Techniques for Skin Cancer Classification*. in *International Conference on Artificial Intelligence and its Applications in the Age of Digital Transformation*. 2025. Springer.
12. Gallazzi, M., et al., *A large dataset to enhance skin cancer classification with transformer-based deep neural networks.* IEEE Access, 2024. **12**: p. 109544-109559.
13. Akter, M., et al., *An Integrated Deep Learning Model for Skin Cancer Detection Using Hybrid Feature Fusion Technique: M. Akter et al.* Biomedical Materials & Devices, 2025. **3**(2): p. 1433-1447.
14. Pukhta, I.R. and R.K. Rout, *Hybrid optimization assisted deep ensemble classification framework for skin cancer detection.* Multimedia Tools and Applications, 2025. **84**(1): p. 63-94.
15. Vijayalakshmi, M., *Melanoma skin cancer detection using image processing and machine learning.* International Journal of Trend in Scientific Research and Development (IJTSRD), 2019. **3**(4): p. 780-784.
16. Salian, S.R. and S.D. Sawarkar, *Melanoma Skin Lesion Classification Using Improved Efficientnetb3.* Jordanian Journal of Computers and Information Technology, 2022. **8**(1).
17. Gouda, W., et al. *Detection of skin cancer based on skin lesion images using deep learning*. in *Healthcare*. 2022. MDPI.
18. Fraiwan, M. and E. Faouri, *On the Automatic Detection and Classification of Skin Cancer Using Deep Transfer Learning.* Sensors, 2022. **22**(13): p. 4963.