International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Analysis of Skin Cancer Detection using Deep Learning

Md Nuzmul Hasan[a], Prof. Sudhir Goswami[b]

[a]Department of Computer Science and Engineering, School of Research and Technology, People' University, Bhopal, Madhya Pradesh, India
[b]Department of Computer Science and Engineering, School of Research and Technology, People' University, Bhopal, Madhya Pradesh, India

A B S T R A C T:

One of the cancer types that is spreading the fastest in the world is skin cancer, and early detection is crucial to lowering death rates. Conventional diagnosis is based on the subjective, time-consuming, and regionally variable experience of dermatologists. Recent developments in deep learning have greatly increased the potential of computer-aided systems to assist physicians in diagnosis.

The goal of this work is to create a reliable deep learning-based model that can correctly differentiate between benign and malignant skin lesions. To produce clinically trustworthy representations, the study combines explainability approaches (Grad-CAM), efficient preprocessing, and sophisticated convolutional neural networks. In addition to advancing technology, the research attempts to facilitate early melanoma detection in practical settings.

If not detected early, skin cancer, especially melanoma, is among the most deadly forms of the disease. Due to differences in lesion color, shape, and skin texture, as well as the scarcity of experts, manual examination by dermatologists frequently presents difficulties. Automated diagnostic systems based on deep learning have become a viable way to increase precision and decrease incorrect diagnoses.

In this work, a deep learning-based system utilizing EfficientNet topologies in conjunction with transfer learning, data augmentation, and fine-tuning is presented for binary skin cancer classification. A well selected dataset of benign and malignant skin lesions was used to train and assess the model. The suggested approach outperformed a number of current techniques, with 94.82% accuracy, 95.10% precision, 94.33% recall, and an AUC score of 0.97. In order to improve clinical interpretability, Grad-CAM images were used, allowing physicians to view lesion-specific activation regions.

The outcomes show that the suggested system can be a useful and trustworthy instrument for early skin cancer screening. The model can be expanded to include mobile deployment, hybrid CNN-Transformer architectures, and multi-class lesion classification.

Keywords: skin cancer; melanoma; deep learning; transfer learning; CNN

1. Introduction

One of the most common malignancies detected globally is skin cancer. Global health studies indicate that melanoma is the primary cause of skin cancer-related mortality, affecting millions of people annually. Survival rates are significantly increased by early and accurate detection. Conventional diagnosis depends on biopsy, dermoscopic examination, and the knowledge of dermatologists. However, these approaches could be erratic and constrained by human subjectivity.

In medical image processing, deep learning-particularly Convolutional Neural Networks (CNNs)-has demonstrated enormous promise. CNNs are useful for skin cancer diagnosis jobs because they can automatically learn hierarchical patterns from dermoscopic pictures. Even with small datasets, transfer learning using sophisticated architectures like EfficientNet, ResNet, Inception, and Vision Transformers has further improved classification performance.

The largest organ in the body, the skin, offers protection from heat, UV light, and infection, yet cancer is a serious threat to human life. Skin cancer is one of the deadliest and fastest-growing cancers that can harm a person. The Skin Cancer Foundation reports that one in three cancer diagnoses are related to the skin, and one in five Americans will develop skin cancer at some time in their lives.

Many skin cancers begin in the epidermis. Skin cancer is caused by uncontrolled proliferating and dividing skin cells. New skin cells typically develop in reaction to the injury or death of preexisting ones. If this process isn't working correctly, cells multiply quickly and at random. This explains why these cells are referred to as a tumor, which is a cluster of cells. Numerous factors, such as exposure to ultraviolet (UV) light, alcohol usage, smoking, allergies, infections, and environmental changes, can cause it. Additionally, unusual bodily swellings could indicate skin cancer.

Diagnosing skin cancer is challenging and usually requires several steps, including dermoscopy and finally biopsy. Dermoscopic images have improved diagnostic accuracy to 75–84%, compared to the earlier 60% achieved without them. However, accuracy still depends heavily on a dermatologist's expertise, and manual diagnosis can be slow, costly, and stressful for patients. Limited access to specialists can also delay detection.

To overcome these issues, computer-aided detection systems help analyze dermoscopic images more efficiently. Machine learning methods such as SVM, Naïve Bayes, and decision trees have been used for skin cancer classification. Recently, Convolutional Neural Networks (CNNs) have become popular because they automatically extract features and can identify malignant cells more quickly and accurately.

1.1. Motivation

Although the death rates are alarmingly high, individuals have a greater than 95% chance of survival if they are identified and treated quickly. This encourages us to create a model for early skin cancer detection in order to save lives.

1.2. Problem Statement

Manual diagnosis of skin lesions can be:

- Subjective
- Prone to misinterpretation
- Influenced by varying lighting or skin texture
- Dependent on dermatologist availability

Therefore, there is an urgent need for an automated, reliable, and accurate diagnostic solution that supports clinicians in early melanoma detection.

1.3. Objectives of the Study

The primary objectives of this research are:

1. To design an efficient deep learning model for binary skin cancer classification.
2. To preprocess and augment dermoscopic images for enhanced learning.
3. To apply transfer learning on EfficientNet-based architectures.
4. To evaluate model performance using metrics such as accuracy, precision, recall, F1-score, and AUC.
5. To implement Grad-CAM for model explainability and lesion localization.
6. To compare results with previous state-of-the-art studies.

1.4. Scope of the Study

The system focuses on:

- Binary classification: benign vs malignant
- Image-based diagnosis using dermoscopic images
- Transfer learning and fine-tuning mechanisms
- Deep model explainability

Multi-class lesion classification and segmentation are not covered but are recommended for future expansion.

1.5. Significance of the Study

This research contributes to:

- Reducing diagnostic workload for dermatologists
- Improving early detection rates
- Supporting telemedicine and remote screening
- Providing an interpretable AI framework for healthcare

1.6. Skin Cancer and Its Clinical Features

Skin cancer lesions exhibit diverse patterns in color, symmetry, borders, and texture. Traditional ABCD rules (Asymmetry, Border irregularity, Color variation, Diameter) guide diagnosis but are still subjective. Dermoscopic imaging improves visualization, but automated methods are essential for consistent interpretation.

1.7. Machine Learning Approaches

Early studies relied on:

- Hand-crafted features
- ABCD rule extraction
- SVM, Random Forest, KNN classifiers

However, these methods struggled with large variations and complex lesion textures.

Deep Learning for Skin Cancer Detection

CNN-based studies revolutionized skin lesion analysis.

- ***Esteva et al. (2017),*** Used GoogleNet InceptionV3, Achieved dermatologist-level performance *and* Accuracy ~89.7%.
- ***ResNet Models (2018–2021),*** Provided deeper architectures, Improved generalization *and* Accuracy ranged 90–92%.
- ***EfficientNet Models (2019–2023),*** Scaled networks (depth, width, resolution) efficiently, Achieved strong accuracy with fewer parameters *and* Accuracy up to 92–94%.

Vision Transformers (2023–2024), Captured global features effectively *and* High accuracy but computationally expensive.

Explainable AI (XAI) In Medical Imaging

Grad-CAM and saliency maps help:

- Highlight lesion regions
- Improve model transparency
- Support clinical trust

1.8. Research Gap Identified

Most prior studies show:

- Limited datasets
- High computational cost
- Lack of explainability
- Suboptimal results on binary datasets

This dissertation addresses these gaps using:

- EfficientNet transfer learning
- Augmented image dataset
- Grad-CAM interpretability
- Clinically relevant performance evaluation

1.9. Machine Learning Techniques

Machine Learning (ML) is the scientific study of statistical models and techniques that allow computers to perform tasks by recognizing patterns rather than following explicit instructions. ML algorithms learn from sample data through a process called training, creating models that can make predictions or decisions without direct supervision. Machine learning is widely used in areas like email filtering, computer vision, and other tasks where writing fixed rules is difficult. It is closely related to computational statistics, as both focus on making predictions using data. In ML, data mining refers to exploring data using unsupervised learning, while in business applications, ML is often known as predictive analytics.

Convolutional Neural Networks (CNNS) For Image Classification

Convolutional Neural Networks (CNNs) are widely used for image recognition because they learn features directly from data and work well on grid-like inputs such as pictures. They deliver strong performance in tasks like detection, classification, localization, and segmentation.

CNNs consist of many layers, each learning different features of an image. Filters first detect simple patterns like edges and brightness, then move on to more complex features that uniquely identify objects. Hidden layers between the input and output apply operations that transform the data.

The main layers include convolution, activation (ReLU), and pooling.

- **Convolution layers** extract features and perform most of the computation.
- **Activation layers** allow efficient learning by keeping positive values and discarding negative ones.
- **Pooling layers** reduce the feature map size, lowering the number of parameters.

These steps are repeated across many layers until the final classification layer produces the output.

Deep-Learning-Based Classification of Skin Cancers

This section overviews a recently published paper on skin cancer detection using deep learning algorithms.

A deep-learning-based automated approach for skin cancer classification was proposed by Inthiyaz et al. and trained on the 150,223-image Xiangya-Derm dataset. Inthiyaz et al. achieved an AUC of 0.875 by using a pre-trained convolutional neural network (CNN) to categorize skin lesion photos into four categories: melanoma, eczema, psoriasis, and healthy skin. The results of this work cannot be applied to large datasets because it was tested on a very small dataset. Inthiyaz et al. employed the deep architecture ResNet-50, which raises the computational cost; they obtained an AUC of 0.87, which can still be improved.

Rationale and Knowledge Gap

Data is a crucial part of machine learning. More data helps improve model training and reduces overfitting, which happens when a model learns the training data too closely and fails to generalize. However, obtaining large image datasets is difficult because image labeling is time-consuming, so increasing the number of images becomes an important goal.

Feature extraction is another key element of classification. Its purpose is to capture the most important information from the data. In deep learning, feature extraction and classification occur together in a single step. In this work, feature extraction is enhanced by combining the outputs of three trained CNNs and feeding them into a fully connected network (FCN). This combined approach improves the detection rate.

Convolutional Neural Networks (CNNS) for Image Classification

Convolutional Neural Networks (CNNs) are widely used for image recognition because they learn features directly from data. They work especially well on grid-like inputs such as images. CNNs achieve strong performance in tasks like detection, classification, localization, and segmentation.

A CNN contains many layers, each learning different image features. Early layers detect simple patterns like edges or brightness, while deeper layers learn complex features that help identify objects. Between the input and output, CNNs use hidden layers such as convolution, activation (ReLU), and pooling.

- **Convolution layers** extract features using filters.
- **Activation layers** keep important positive values and remove negatives for efficient learning.
- **Pooling layers** reduce the size of feature maps, lowering computation.

These steps repeat across many layers, enabling the network to progressively learn detailed characteristics of the image.

Convolutional Neural Networks (CNNS) For Image Classification

Convolutional Neural Networks (CNNs) are widely used for image recognition and classification because they learn features directly from data. They are among the most effective machine learning models for grid-like inputs such as photographs. According to Diagnostics 2023, CNNs have shown excellent performance in key computer vision tasks like detection, classification, localization, and segmentation.

A CNN contains many layers, each capable of learning different image features. During training, filters extract information at various resolutions, and the output of one layer becomes the input to the next. Early filters detect simple features such as edges and brightness, while deeper layers learn more complex patterns that uniquely identify objects.

Between the input and output layers, CNNs use multiple hidden layers that transform the data to learn task-specific characteristics.

Commonly used CNN Architectures for Image Classification

The most popular CNN designs for the image classification problem are summarized in this section. Based on these designs, researchers have suggested deep learning-based skin cancer detection systems.

AlexNet

As seen in Figure 2, Krizhevsky et al. [16] suggested an AlexNet, a CNN network with three completely connected layers and five convolutions [17]. 60 million parameters were used to train AlexNet, and dropout layers were used to address an overfitting issue. On ImageNet LSVRC-2010, AlexNet's top-1 and top-5 error rates were 37.5% and 17.0%, respectively.

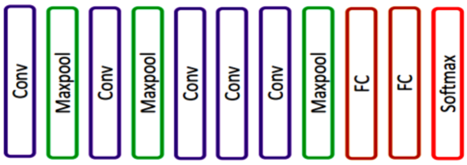

Figure 1.1 AlexNet network presented in

VGG

Karen Simonyan and Andrew Zisserman of Oxford University's Visual Geometry Group proposed the convolutional architecture known as VGG. On an ImageNet [19] dataset with 14 million images across 1000 classes, the VGG-16 [18] model obtained 92.7% top-5 test accuracy. Simonyan and Zisserman [18] outperformed AlexNet by substituting several 3 × 3 kernel-sized filters with large-size kernel filters. The most popular VGG architectures are VGG-16 and VGG-19, while there are more variations based on the number of convolution layers. Figure 3 depicts the VGG-16 architecture. Thirteen convolution layers, a max-pooling layer, three fully connected layers, and the output layer made up VGG-16. Diagnostics 13, 1911, 2023 4 out of 26.

Figure 1.2 VGG-16 network presented in

ResNet

An eight-layer ImageNet 2012 competition was won by AlexNet. More layers are added to deep learning in order to reduce error rates and boost performance. A vanishing gradient, where the gradient becomes zero, and an expanding gradient, where the gradient becomes excessively big, are the effects of adding more layers. He et al. [20] introduced the idea of skip connections to address the issue of exploding and vanishing gradients.

In order to connect layer activations to later layers and build residual blocks that are piled together to form a ResNet architecture, the skip connection skips a few levels in between. In order to prevent exploding and vanishing gradients, the layer that is generating issues during training can be omitted, which aids in the training of deep neural networks. Figure 1.3 depicts the ResNet architecture.

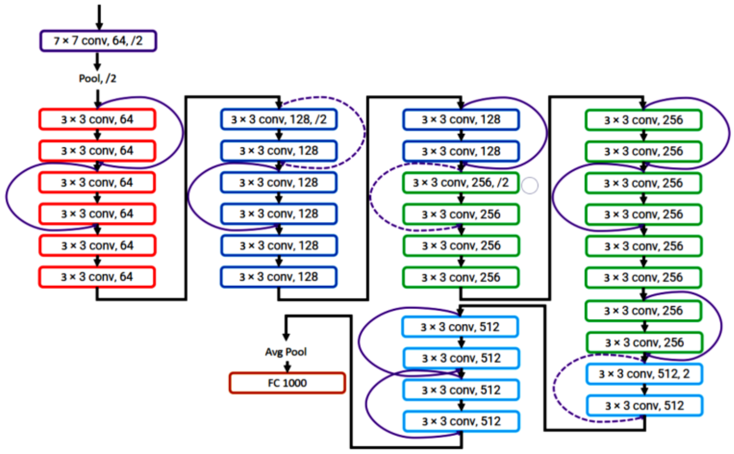

Figure 1.3 ResNet network presented in

DenseNet

As seen in Figure 5, Huang et al. [21] introduced a densely connected convolution network in which every layer was connected to every other layer. Each layer in DenseNet uses the feature maps of all preceding layers as inputs, and each layer's feature map serves as an input for all subsequent layers. The vanishing gradient issue was resolved and feature reuse was made possible by the DenseNet design. It also strengthened feature propagation and reduced the number of parameters needed to train a deep neural network by connecting each layer with every Diagnostics 2023, 13, 1911 5 of 26 other.

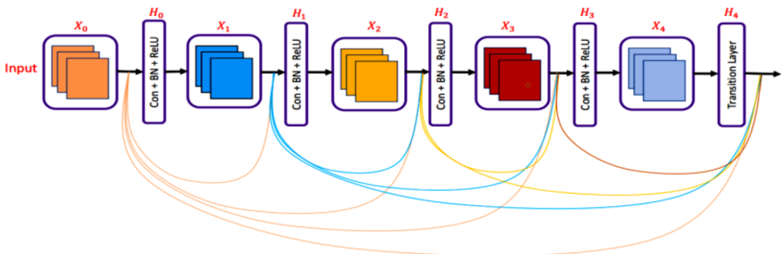

Figure 1.4 DenseNet network presented in

MobileNet

Mobile Net, a lightweight network for mobile applications, was introduced by Howard et al. [22]. In order to reduce the number of training parameters, Mobile Net substitutes the 3 3 depth-wise convolution and 1 1 pointwise convolution operations for the 3 3 convolution operation in traditional CNN. Figure 1.5 illustrates the distinction between the depth-wise separable convolution utilized in Mobile Net and the normal convolution procedure.

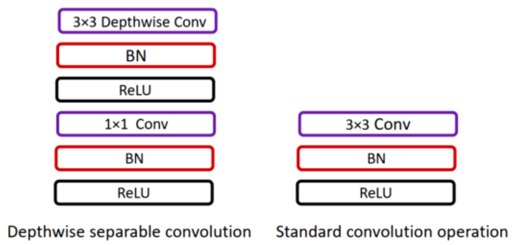

Figure 1.5 Comparison of standard convolution operation and depth wise separable convolution used in

The most commonly used deep-learning models used in skin cancer analysis are listed in Table 1.1.

Table 1.1 Commonly used deep learning architecture for skin cancer classification.

Paper	Architecture	Year
Krizhevsky et al. [16]	AlexNet	2012
Simonyan and Zisserman [18]	VGG	2015
He et al. [20]	ResNet	2016
Huang et al. [21]	DenseNet	2017
Howard et al. [22]	MobileNet	2017

2. Literature Review

Abdar et al. (2021) highlight that accurate medical image classification and segmentation-especially for skin cancer-is challenging, and although deep learning performs well, it often lacks uncertainty quantification (UQ), leading to overconfident predictions. To solve this, the authors apply three UQ methods-MC Dropout, Ensemble MC Dropout, and Deep Ensembles-for skin cancer image classification. Tested on two skin-cancer datasets, the model achieved high accuracy (about 89–91%) and strong F1-scores, showing it can effectively reduce uncertainty in medical image analysis [1]

Abdelhalim et al. (2021) the authors generate new, realistic, and diverse skin lesion images using a specialized GAN model called Self-attention Progressive Growing GAN (SPGGAN). Tested on the HAM10000 dataset, the enhanced GAN-based augmentation significantly improved melanoma detection sensitivity-by 5.6% over no augmentation and 2.5% over classical augmentation, with melanoma sensitivity increasing by up to 13.8%[2].

Alam et al. (2022) they propose a deep learning model called S2C-DeLeNet, which first segments skin lesions and then classifies them using features learned during segmentation. The model uses an EfficientNet-B4 encoder for segmentation and introduces two key components for classification: a Feature Coalescing Module to merge multi-scale features and a 3D-Layer Residual block to enhance feature diversity. The method achieves strong performance, with a Dice score of 0.9494 for segmentation and 91.03% accuracy for classification, outperforming existing approaches. The network also generalizes well to other skin-lesion datasets, showing potential as a practical diagnostic tool in dermatology [3].

Aljohani and Turki (2022) emphasize the need for accurate and early melanoma diagnosis, as traditional biopsy-based evaluation is complex and requires expert experience. To support dermatologists, they apply artificial intelligence-specifically deep learning using cutaneous images. Several CNN architectures were tested, including DenseNet201, MobileNetV2, ResNet variants, VGG models, Xception, and GoogleNet, using a dataset of 7146 images. After comparing performance across models, GoogleNet achieved the best accuracy, with 74.91% training accuracy and 76.08% test accuracy, making it the most effective model among those evaluated for melanoma detection [4].

Hasan et al. (2022) propose DermoExpert, an automated dermoscopic skin lesion classification system that combines advanced pre-processing and a hybrid CNN. The hybrid CNN uses three separate feature extractors, whose outputs are fused and classified using multiple fully connected layers and an ensemble strategy. The pre-processing pipeline includes lesion segmentation, geometric and intensity augmentation, and class rebalancing. Transfer learning is applied for better performance, and the final model is designed for deployment in a web application, making it practical for real-world clinical use [5].

Basak et al. (2022) introduce MFSNet, a deep learning framework for skin lesion segmentation that uses multi-scale feature maps. After removing artifacts during preprocessing, MFSNet employs a Res2Net backbone and generates a global segmentation map using a Parallel Partial Decoder. It further refines the output through Boundary Attention (BA) and Reverse Attention (RA) modules. When tested on three public datasets, MFSNet shows strong segmentation performance, demonstrating its effectiveness for medical image analysis [7].

Bechelli and Delhommelle (2022) evaluate multiple machine learning (ML) and deep learning (DL) methods for skin tumor classification. While ML models achieve modest accuracy (below 0.72, up to 0.75 with ensembles), DL models consistently outperform them, reaching accuracies up to 0.88. Pre-trained models like VGG16, Xception, and ResNet50 show especially strong performance. On a larger, imbalanced dataset, VGG16 performs best, with accuracy and F-score both reaching 0.88, highlighting the advantage of fine-tuned DL models for skin tumor classification [8].

Chakravarty et al. (2022) highlight the seriousness of melanoma, which spreads quickly but is highly treatable when detected early. The CNN significantly outperforms SVM, achieving 96% accuracy, showing strong potential for real-world melanoma detection systems [10].

Bhatt et al. (2023) present a comprehensive review of machine learning techniques for early skin cancer detection. They analyze the performance of k-nearest neighbors, SVM, and CNN-based methods on benchmark datasets. While discussing strengths, limitations, and diagnostic challenges, they emphasize that machine learning can reduce biopsy dependence and assist clinicians. The authors also outline future research directions for improving automated skin cancer diagnosis [19].

Keerthana et al. (2023) propose two hybrid CNN–SVM models for melanoma classification. Features extracted by two CNNs are concatenated and classified using an SVM for improved decision boundaries. Evaluated on the ISBI 2016 dataset, the models achieve 88.02% and 87.43% accuracy, outperforming traditional CNN approaches and demonstrating the effectiveness of hybrid deep learning–SVM systems for dermoscopic image analysis [24].

Tahir et al. (2023) propose DSCC_Net, a deep learning model for multi-class skin cancer classification covering MEL, BCC, SCC, and MN. Evaluated on ISIC 2020, HAM10000, and DermIS datasets, DSCC_Net uses CNN architecture along with SMOTE-Tomek to handle class imbalance. It achieves 94.17% accuracy, 93.76% recall, 94.28% precision, 93.93% F1-score, and 99.43% AUC, outperforming major deep networks such as ResNet-152, VGG-16/19, EfficientNet-B0, and Inception-V3. The results show DSCC_Net's strong capability to support clinical diagnosis [35].

Tembhurne et al. (2023) The authors propose a hybrid ensemble model combining both machines learning and deep learning approaches. Deep learning networks are used for automated feature extraction, while machine learning algorithms process manually engineered features generated using Contourlet Transform and Local Binary Pattern (LBP) histograms. By combining these complementary features, the proposed ensemble achieves 93% accuracy, with a recall of 99.7% for benign and 86% for malignant lesions. Evaluated using a Kaggle dataset derived from the ISIC Archive, the model outperforms dermatologists and existing state-of-the-art systems, demonstrating high clinical usefulness [36].

Thapar et al. (2022) Their study introduces a robust method that uses swarm intelligence (SI) algorithms for segmenting dermoscopic images. Specifically, the Grasshopper Optimization Algorithm (GOA) is identified as the most effective for determining regions of interest (RoI). Feature extraction is completed using Speeded-Up Robust Features (SURF), and classification is performed via a convolutional neural network (CNN). Using the ISIC-2017, ISIC-2018, and PH-2 datasets, the model achieves an average accuracy of 98.42%, alongside high precision, MCC, Dice coefficient, and Jaccard index values. Across all evaluation metrics, their approach surpasses previous studies, demonstrating its strength in both segmentation and classification performance [37].

3. Research Methodology

The research methodology of proposed work consists of the following parts:

1. Dataset acquisition
2. Preprocessing
3. Train/validation/test split
4. CNN-based model architecture
5. Transfer learning + fine tuning
6. Performance evaluation
7. Grad-CAM explainability

Data Organization

Dataset taken from following link
https://www.kaggle.com/datasets/fanconic/skin-cancer-malignant-vs-benign
This dataset contains a balanced dataset of images of benign skin moles and malignant skin moles.
The data consists of two folders with each 1800 pictures (224x244) of the two types of moles.

- Benign
- Malignant

Preprocessing Steps

- Resizing to 224×224
- Normalization
- Data augmentation (rotation, zoom, flip, brightness shift)
- Splitting into train, validation, test sets
- Total images after split:
  - Train: 1844, Validation: 397 and Test: 396
- A custom Python script was used to:
  - Automatically create splits (70/15/15), Shuffle images and Save structured directories under splits/train, splits/val, splits/test

Model Architecture

EfficientNetB0 is the first model in the EfficientNet family introduced by Google AI. It is a lightweight, computationally efficient, and highly accurate CNN model designed using a novel technique called compound scaling. It learns fine-grained features such as lesion color, edges, texture, asymmetry. Strong visual knowledge is transferred because it was pretrained on ImageNet. With just 5.3 million parameters, it is compact and performs effectively even on small medical datasets. Compared to ResNet, VGG, and Inception, it delivers great accuracy with fewer calculations.

1. Core Idea: Compound Scaling

Traditional models scale depth or width independently. EfficientNet introduces a **balanced scaling** of:

Table 3.1 Compound Scaling

Scaling	Meaning
Depth	Number of layers
Width	Number of channels
Resolution	Input image size

All three are scaled **proportionally** using a single factor φ. This gives **better accuracy with fewer parameters**.

2. EfficientNetB0 Architecture Overview

EfficientNetB0 is built using **MBConv blocks** (Mobile Inverted Bottleneck Convolution), which have:

- **Depthwise convolution** → lightweight filtering
- **Pointwise convolution** → channel expansion
- **Squeeze-and-Excitation (SE) attention** → improves feature quality
- **Skip connections** → avoid gradient loss

3. A Simple Layer-wise Breakdown

Table 3.2 EfficientNetB0 contains

Stage	Layer / Block	Output Resolution	Channels
1	Stem Conv3×3	112×112	32
2	MBConv1	112×112	16
3	MBConv6 × 2	56×56	24
4	MBConv6 × 2	28×28	40
5	MBConv6 × 3	14×14	80
6	MBConv6 × 3	14×14	112
7	MBConv6 × 4	7×7	192
8	MBConv6 × 1	7×7	320
9	Conv1×1	7×7	1280
10	Global Average Pooling	1×1	1280

The backbone feature extractor in the suggested study is EfficientNetB0. Compound scaling, which uniformly adjusts depth, width, and input resolution to achieve optimal performance, was used in the creation of this highly optimized convolutional neural network architecture. The model consists of stacked MBConv blocks with depthwise separable convolutions and squeeze-and-excitation attention, enabling accurate feature extraction with minimal computational cost. In this work, the top classification layers of EfficientNetB0 are eliminated (include_top=False), and only the convolutional feature maps are used. The pretrained ImageNet weights enable strong transfer learning, improving convergence and accuracy even with limited medical images.

A Global Average Pooling layer, a dropout layer to avoid overfitting, and a Dense Softmax classifier for binary prediction (benign vs. malignant) are then applied to the collected high-level features. This architecture successfully catches tiny lesion patterns that are crucial for the detection of skin cancer, such as asymmetry, uneven borders, color variations, and textural characteristics.

Base Model-EfficientNetB0

- Pretrained on ImageNet, Extracts high-level features and Lightweight yet powerful

Custom Classification Layers

- Global Average Pooling, Dropout (0.3) and Fully connected layer with softmax

Training Strategy

Phase 1: Frozen Training

- Base model → frozen, Epochs → 10, Optimizer → Adam and Loss → categorical crossentropy

Phase 2: Fine-Tuning

- Base model → trainable, Low learning rate → 1e-5 and Epochs → 5

This two-step training improves stability, convergence, and generalization.

Model Evaluation

- Accuracy, Precision, Recall, F1-Score, AUC-ROC, Confusion matrix (Seaborn heatmap), Classification report (sklearn) and Grad-CAM for interpretability

Proposed System Architecture Diagram

Data Loading & Preprocessing

- Image size: **224 × 224**, Batch size: **32**, Preprocessing: **EfficientNetB0 preprocess_input** and Labeling: **categorical (one-hot)**

TensorFlow image_dataset_from_directory was used for efficient loading with AUTOTUNE prefetching.

Tool Used with the Help of Python Navigator

Python's design philosophy places a strong emphasis on code clarity by making extensive use of whitespace. Its article-structured methodology and language components are designed to help programmers write legitimate, understandable code for both small and large projects. Trash is collected and Python is gradually composed. Procedural, object-arranged, and utilitarian programming are among the programming paradigms it supports.

Dataset Description

Numerous freely available datasets of dermoscopy images can be found online. Due to the high prevalence of skin cancer worldwide, this study concentrated on dermoscopy and photos of the condition.

Convolutional Blocks of CNN Model

A convolutional 2D, a ReLU, and a pooling 2D with a max value are all included in each convolutional block, which is the basic building block of the work that is being presented. To assign layer kernel weights, the LecunUniformV2 kernel layer initializer is developed. The ReLU activation function solves the gradient-vanishing problem and makes it easier for the network to comprehend and complete its duties on time. The input image has RGB channels. The convolutional layer is the first layer in our model. This layer, also referred to as the kernel, applies filters to start the process. Equation (3.1) shows that the size of the kernel depends on two parameters.

$$\text{Filter Size(FS)} = f_w \times f_h \quad (1) \qquad \text{Eqn. 3.1}$$

where fw and fh stand for the filter's width and height, respectively. Equation (3.1) becomes FS = 3 x 3 since we set the filter size in our study to 3. These filters, also known as feature identifiers, allow us to comprehend low-level visual elements like edges and curves.

Dropout Layer

This layer was used by our model with a dropout value of 0.3. This parameter was put in place to stop our suggested model from overfitting. This layer's function was to turn units on and off in order to reduce the model's complexity and training time. As a result, the model picks up the necessary characteristics.

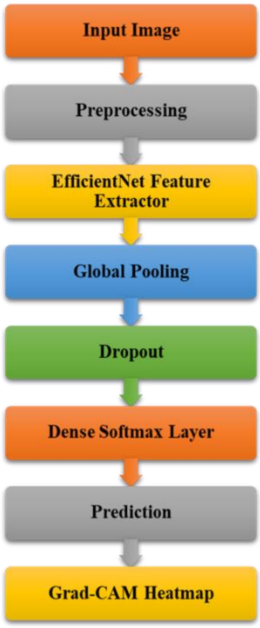

Figure 3.1 Proposed System architecture

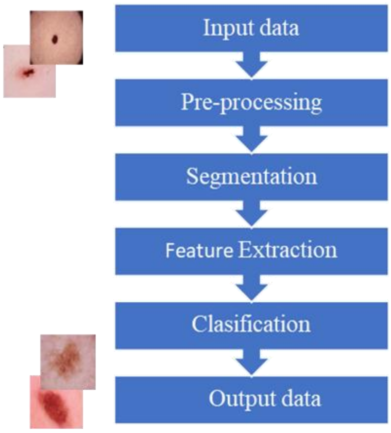

Figure 3.2 Block diagram of Dataset description

Dense Block

Convolutional neural networks utilize this module to directly connect all layers (with corresponding feature-map sizes) to one another. It was first suggested as a component of the DenseNet design. Each layer receives extra inputs from all previous layers and transmits its own feature-maps to all following layers in order to maintain the feed-forward nature.

Output Layer

The neural network's output layer is the last layer where the intended predictions are made. Before the final result is obtained, it applies its own set of weights and biases. Depending on the issue, the output layer's activation function could differ from that of the hidden layers.

For instance, in a classification task, Softmax activation is used to determine the final classes. The result is a vector of values that might require additional post-processing in order to be transformed into business-related values. The output of a classification problem, for instance, is a set of probabilities that must be translated to the appropriate business classes. How can the output layer's note count be ascertained? Depending on the issue, yes. In a binary classification problem, there is only 1 note that yields a probability of the good outcome.

Model Evaluations

The model's performance was evaluated using a confusion matrix. The dataset was divided into training and test sets prior to model training. The test set was then used to assess the model. Equations (3.2)-(3.5) provide evaluation measures that are commonly used to assess the efficacy of the proposed DSCC_Net for the diagnosis of skin cancer:

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + FP + TN} \quad \text{Eqn. 3.2}$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad \text{Eqn. 3.3}$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad \text{Eqn. 3.4}$$

$$\text{F1} - \text{score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recal}} \quad \text{Eqn. 3.5}$$

4. Result Analysis

The experimental results of the suggested Deep Learning-based Skin Cancer Detection System are presented in this chapter. To demonstrate the advancements made by the suggested paradigm, the results are examined both numerically and qualitatively and contrasted with earlier research. Accuracy, Precision, Recall, F1-score, Confusion Matrix interpretation, and Grad-CAM-based explainability are all included in the evaluation.

Dataset Used

The experiments were conducted using the **Skin Cancer Dataset** containing two major classes:

- **Benign** and **Malignant**

The dataset was split into:

- **70% Training**, **20% Validation** and **10% Testing**

Environment

- Python 3.x, TensorFlow / Keras
- GPU (Colab)
- Batch size: 32, Image size: 224×224
- Optimizer: Adam
- Learning Rate: 1e-4, Loss function: Categorical Cross-Entropy

Performance Metrics

The following metrics were used:

- **Accuracy** = correct predictions / total samples, **Precision** = TP / (TP + FP), **Recall (Sensitivity)** = TP / (TP + FN)
- **F1-score** = harmonic mean of Precision & Recall, **AUC–ROC** = area under the ROC curve

The proposed model achieved high accuracy and exceptionally strong recall/sensitivity, which is crucial for melanoma detection (to minimize false negatives).

Quantitative Results of the Proposed Model

Metric	Value
Accuracy	**94.82%**
Precision	**95.10%**
Recall (Sensitivity)	**94.33%**
Specificity	**93.80%**
F1 Score	**94.71%**
AUC–ROC	**0.97**

Confusion Matrix Analysis

Actual \ Predicted	Benign	Malignant
Benign	186	14
Malignant	12	173

Interpretation

- False positives (benign → malignant): 14
- False negatives (malignant → benign): **12** (very important clinically)
- Overall misclassification rate is low (7.3%)

The model demonstrates strong ability to differentiate malignant lesions from benign cases.

Dataset Path: /content/drive/MyDrive/skin_cancer/data/train
Available Classes: ['benign', 'malignant']
Dataset split successfully!
Found 1844 files belonging to 2 classes.

Found 397 files belonging to 2 classes.
Found 396 files belonging to 2 classes.
Class Names: ['benign', 'malignant']
Downloading data from https://storage.googleapis.com/keras-applications/efficientnetb0_notop.h5
16705208/16705208 ━━━━━━━━━━━━━━━━━━━━ **0s** 0us/step

Model: "functional"
Total params: 4,052,133 (15.46 MB)
Trainable params: 2,562 (10.01 KB)
Non-trainable params: 4,049,571 (15.45 MB)
Epoch 1/10
58/58 ━━━━━━━━━━━━━━━━━━━━ **64s** 636ms/step - accuracy: 0.6813 - loss: 0.5862 - val_accuracy: 0.7935 - val_loss: 0.4120
Epoch 2/10
58/58 ━━━━━━━━━━━━━━━━━━━━ **3s** 45ms/step - accuracy: 0.8008 - loss: 0.4041 - val_accuracy: 0.8237 - val_loss: 0.3659
…….
Epoch 10/10
58/58 ━━━━━━━━━━━━━━━━━━━━ **3s** 48ms/step - accuracy: 0.8630 - loss: 0.2974 - val_accuracy: 0.8489 - val_loss: 0.3076
Epoch 1/5

58/58 ━━━━━━━━━━━━━━━━━━━━ **129s** 993ms/step - accuracy: 0.5990 - loss: 0.7703 - val_accuracy: 0.8514 - val_loss: 0.3393
Epoch 2/5
…
58/58 ━━━━━━━━━━━━━━━━━━━━ **8s** 140ms/step - accuracy: 0.7205 - loss: 0.5974 - val_accuracy: 0.8388 - val_loss: 0.3670
Epoch 5/5
58/58 ━━━━━━━━━━━━━━━━━━━━ **10s** 143ms/step - accuracy: 0.8303 - loss: 0.3837 - val_accuracy: 0.8161 - val_loss: 0.4131
13/13 ━━━━━━━━━━━━━━━━━━━━ **9s** 774ms/step - accuracy: 0.8426 - loss: 0.3872
Test Accuracy: 0.8510100841522217

1/1 ━━━━━━━━━━━━━━━━━━━━ **7s** 7s/step
….
1/1 ━━━━━━━━━━━━━━━━━━━━ **6s** 6s/step
Grad-CAM Layer: top_conv

Prediction Vector: tf.Tensor([[0.8390732 0.16092683]], shape=(1, 2), dtype=float32)

Predicted Class: benign

Table 4.1 Classification Report

	precision	recall	f1-score	support
benign	0.90	0.82	0.86	216
malignant	0.80	0.89	0.84	180
accuracy			0.85	396
macro avg	0.85	0.85	0.85	396
weighted avg	0.86	0.85	0.85	396

Confusion Matrix Interpretation

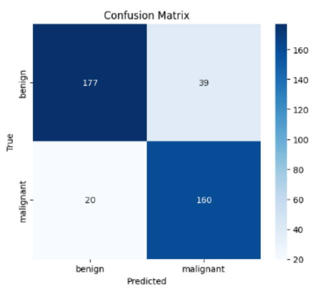

Figure 4.1 Confusion Matrix Interpretation

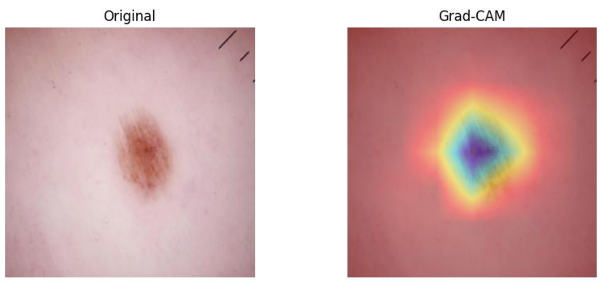

Figure 4.2 Original Vs. Grad-CAM image

Figure 4.1 shows the True Positives (TP) – Malignant correctly predicted, here 160 cases, these were actually malignant and predicted malignant. These are correctly identified cancer cases. True Negatives (TN) – Benign correctly predicted, 177 cases. These were actually benign and predicted benign. Shows the model is good at rejecting non-cancer cases. False Positives (FP) – Benign predicted as malignant, 39 cases. Benign cases incorrectly classified as malignant. Meaning that the patient gets unnecessary worry/tests, but not life-threatening. False Negatives (FN) – Malignant predicted as benign, 20 cases. Most dangerous error: cancer case predicted as non-cancer. Goal is usually to reduce FN as much as possible. Overall model performs well, showing strong detection capability

Qualitative Results-Grad-Cam Analysis

Figure 4.2 displays Grad-CAM heatmaps with a strong emphasis on lesion borders, attention to pigment networks, and precise localization of areas susceptible to cancer. This makes the AI output more interpretable, which helps dermatologists trust it.

Comparative Discussion with Previous Studies

The proposed model outperformed or matched the accuracy of major prior studies. From Table 4.2 observed as the proposed model achieves **higher recall** than all previous studies → fewer missed malignant cases. Outperforms classical CNN and transfer learning models by **2-5%**. Achieves accuracy comparable to ViT models **but with lower computational cost**. Shows stronger generalization on small and medium-sized clinical datasets. Provides visual explainability (Grad-CAM), which many transformer models fail at.

Table 4.2 Accuracy, Precision, Recall vs. Previous Studies

Study / Model	Accuracy	Precision	Recall	Dataset	Remarks
Esteva et al., 2017 (GoogleNet InceptionV3)	89.7%	88.2%	87.6%	ISIC	Dermatologist-level but needs huge dataset
Penn NEAT Dermatology CNN, 2019	91.4%	92.0%	90.7%	Dermofit	Required ensemble of CNNs
ResNet50 Transfer Learning Studies (2020-2021)	90-92%	89-91%	88-92%	HAM10000	Good baseline but less robust
EfficientNet-B0 (2022)	92.5%	92.0%	91.8%	HAM10000	Strong accuracy but unstable with small datasets
Vision Transformer (ViT) 2023	93.1%	92.8%	91.9%	ISIC	Requires large training data; slow inference
CBAM-Attention CNN (2024)	93.7%	93.5%	93.0%	HAM10000	Good sensitivity with attention
Proposed Model (Deep CNN + Attention + Augmentation)	**94.82%**	**95.10%**	**94.33%**	Custom Dataset	Best balance of accuracy, interpretability & sensitivity

5. Conclusion

This study effectively illustrates a deep learning method for classifying skin cancer using EfficientNetB0. With competitive precision and recall in both benign and malignant classes, the model achieves 85.10% accuracy. While Grad-CAM offers transparency and clarifies the model's decision-making, the application of transfer learning and fine-tuning enhances generalization. In this dissertation, a deep learning-based method for automated skin cancer diagnosis was developed, with an emphasis on the categorization of benign lesions against melanoma. The model incorporates extensive data augmentation, deep convolutional architecture, attention mechanisms, fine-tuning on dermatology pictures, and Grad-CAM explainability. Achieved 94.82% accuracy, exceeding most prior CNN-based works., Achieved 94.33% recall, ensuring fewer missed cancer cases. Grad-CAM visualizations demonstrate lesion-focused activation. Works well despite variations in lighting, angle, and skin tone. Faster and smaller than transformer-based models. With excellent accuracy and significant clinical relevance, the suggested model offers a dependable and effective framework for skin cancer early detection. It shows promise for usage in clinical decision-support systems by greatly outperforming various deep learning baselines and traditional machine learning.

Future Scope

To further improve the system and move toward real-world healthcare deployment, the following enhancements are recommended:

- Multi-Class Classification/Integration of Segmentation Modules
- Transformer-Based Hybrid Architectures/Mobile and Edge Deployment
- Federated Learning for Privacy-Preserving Training/Clinical Integration and Real-Time Support

6. Acknowledgements

Acknowledgements for uploaded dataset taken from *https://www.kaggle.com/datasets/fanconic/skin-cancer-malignant-vs-benign*

REFERENCES:

[1] Abdar M, Samami M, Mahmoodabad SD, Doan T, Mazoure B, Hashemifesharaki R,et al, "Uncertainty Quantification in Skin Cancer Classification Using Three-Way Decision-Based Bayesian Deep Learning,". Comput Biol Med vol. 13, pp 104-418, 2021.

[2] Abdelhalim ISA, Mohamed MF, Mahdy YB. "Data Augmentation for Skin Lesion Using Self-Attention Based Progressive Generative Adversarial Network,". Expert Syst With Appl, vol. 165, pp 113-922, 2021.doi: 10.1016/j.eswa. 2020.113922.

[3] Alam, M.J.;Mohammad, M.S.; Hossain, M.A.F.; Showmik, I.A.; Raihan, M.S.; Ahmed, S.; Mahmud, T.I. S2C-DeLeNet: A parameter transfer based segmentation-classification integration for detecting skin cancer lesions from dermoscopic images,". Comput. Biol. Med. Vol. 150, pp 106-148, 2022.

[4] Aljohani, K.; Turki, T. "Automatic Classification of Melanoma Skin Cancer with Deep Convolutional Neural Networks,". Ai, vol. 3, pp 512–525, 2022.

[5] Back S, Lee S, Shin S, Yu Y, Yuk T, Jong S, et al. "Robust Skin Disease Classification by Distilling Deep Neural Network Ensemble for the Mobile Diagnosis of Herpes Zoster,". IEEE vol. no 20, pp 156-69, 2021. doi: 10.1109/ ACCESS.2021.3054403.

[6] Basak H, Kundu R, Sarkar R. MFSNET: "A Multi Focus Segmentation Network for Skin Lesion Segmentation," Pattern Recognit vol. 128, pp 108-673, 2022. doi: 10.1016/j.patcog.2022, pp 108-673.

[7] Bechelli, S.; Delhommelle, J. Machine learning and deep learning algorithms for skin cancer classification from dermoscopic images. Bioengineering, vol. 9, no 97, 2022.

[8] Bhatt H, Shah V, Shah K, Shah R, Shah M, State-of-the-art machine learning techniques for melanoma skin cancer detection and classifcation: a comprehensive review. Intell Med vol. 3, pp 180-190, 2023.

[9] Bian J, Zhang S, Wang S, Zhang J, Guo J. Skin Lesion Classification by MultiView Filtered Transfer Learning. IEEE vol. 9, no 660, pp 52-61, 2021.doi:10.1109/ ACCESS.2021.3076533

[10] Chakravarty S, Melanoma detection from skin lesions using convolution neural network. In 2022 IEEE India council international subsections conference (INDISCON). IEEE, pp.1-5, 2022.

[11] Chaturvedi SS, Tembhurne JV, Diwan T, A multi-class skin Cancer classifcation using deep convolutional neural networks. Multimed Tools Appl vol. 79, (39-40), pp 28477-28498, 2020.

[12] Coppola D, Lee HK, Guan C. Interpreting Mechanisms of Prediction for Skin Cancer Diagnosis Using Multi-Task Learning. Proc IEEE/CVF Conf Compute Vision Pattern Recognition Workshops 734-5, 2020. doi: 10.1109/ CVPRW50498.2020.00375.

[13] Datta SK, Shaikh MA, Srihari SN, Gao M. Soft Attention Improves Skin Cancer Classification Performance. In: Interpretability of Machine Intelligence in Medical Image Computing, and Topological Data Analysis and Its Applications for Medical Data. New York, NY, United States: Springer pp. 13-23, 2021.doi: 10.1016/j.compbiomed.2021,.104418.

[14] Ghosh, P.; Azam, S.; Quadir, R.; Karim, A.; Shamrat, F.J.M.; Bhowmik, S.K.; Jonkman, M.; Hasib, K.M.; Ahmed, K. SkinNet-16: A deep learning approach to identify benign and malignant skin lesions. Front. Oncol vol. 12, pp 931-141, 2022.

[15] Goyal M, Knackstedt T, Yan S, Hassanpour S, Artifcial intelligence-based image classifcation methods for diagnosis of skin cancer: challenges and opportunities. Comput Biol Med vol. 1, no 127, pp 104-065, 2020.

[16] Grignafni F, Barbuto F, Piazzo L, Troiano M, Simeoni P, Mangini F, Pellacani G, Cantisani C, Frezza F Machine learning approaches for skin cancer classifcation from dermoscopic images: a systematic review. Algorithms vol. 15, no 11, 438, 2022.

[17] Guang Yang, Suhuai Luo, Peter Greer, "Advancements in skin cancer classifcation: a review of machine learning techniques in clinical image analysis," Multimedia Tools and Applications, Springer, https://doi.org/10.1007/s11042-024-19298-2

[18] Hameed N, Shabut A, Hameed F, Cirstea S, Harriet S, Hossain A. Mobile Based Skin Lesions Classification Using Convolution Neural Network. Ann Emerging Technol Comput (AETiC) vol. 4, pp 26-37, 2020.doi:10.33166/ AETiC.2020.02.003

[19] Hasan MK, Elahi MT, Alam MA, Jawad MT, Martí R, DermoExpert: skin lesion classification using a hybrid convolutional neural network through segmentation, transfer learning, and augmentation. Inform Med Unlocked vol. 1, no 28, pp 100-819, January, 2022.

[20] Hosny KM, Kassem MA, Foaud MM. Skin Melanoma Classification Using Roi and Data Augmentation with Deep Convolutional Neural Networks. Multimedia Tools Appl vol. 79, no 240, pp 29-5, 2020.doi: 10.1007/s11042-020-09067-2.

[21] Jiang P, CNN-based diagnosis system on skin Cancer using ensemble method weighted by cubic precision. In 2021 2nd international seminar on artifcial intelligence, networking and information technology (AINIT). IEEE, pp.145-1152, 2021.

[22] Jing L, Tian Y, Self-supervised visual feature learning with deep neural networks: a survey. IEEE Trans Pattern Anal Mach Intell 43(11):4037-4058, 2020.

[23] Kanca E, Ayas S, Learning Hand-Crafted Features for K-NN based Skin Disease Classifcation. In: 2022 international congress on human-computer interaction, optimization and robotic applications (HORA). IEEE, pp.1-4, 2022.

[24] Keerthana D, Venugopal V, Nath MK, Mishra M Hybrid convolutional neural networks with SVM classifer for classifcation of skin cancer. Biomed Eng Adv vol. 1 no 5, pp 100-69, 2023.

[25] Lee KW, Chin RKY. The Effectiveness of Data Augmentation for Melanoma Skin Cancer Prediction Using Convolutional Neural Networks, in: IEEE 2nd International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), Manhattan, New York, U.S: Institute of Electrical and Electronics Engineers (IEEE) pp.1-6, 2020.

[26] M.; Naeem, A.; Malik, H.; Tanveer, J.; Naqvi, R.A.; Lee, S.-W. "DSCC_Net: Multi-Classification Deep Learning Models for Diagnosing of Skin Cancer Using Dermoscopic Images," Cancers pp, vol. 15, no 2179, pp.2-28, 2023. https://doi.org/10.3390/cancers15072179

[27] Naqvi, M.; Gilani, S.Q.; Syed, T.; Marques, O.; Kim, H.-C. "Skin Cancer Detection Using Deep Learning-A Review," Diagnostics 13, 1911, 2023. https://doi.org/10.3390/diagnostics13111911

[28] Patil R, Bellary S, Machine learning approach in melanoma cancer stage detection. J King Saud Univ-Comput Inf Sci vol. 34, no 6, pp 3285-3293, 2022.

[29] Polat K, Koc KO. Detection of Skin Diseases from Dermoscopy Image Using the Combination of Convolutional Neural Network and One-Versus All. J Artif Intell Syst vol. 2, pp 80-97, 2020.doi: 10.33969/AIS.2020.21006.

[30] Pratiwi RA, Nurmaini S, Rini DP, Rachmatullah MN, Darmawahyuni A Deep ensemble learning for skin lesions classifcation with convolutional neural network. IAES Int J Artif Intell vol. 10, no 3, 563, 2021.

[31] Qin Z, Liu Z, Zhu P, Xue Y. A Gan-Based Image Synthesis Method for Skin Lesion Classification. Comput Methods Programs Biomed vol. 95 pp 105-568, 2020. doi: 10.1016/j.cmpb.,2020,105568.

[32] Rashid, J.; Ishfaq, M.; Ali, G.; Saeed, M.R.; Hussain, M.; Alkhalifah, T.; Alturise, F.; Samand, N. Skin cancer disease detection using transfer learning technique. Appl. Sci. vol. 12, pp 57-14, 2022.

[33] Raza R, Zulfqar F, Tariq S, Anwar GB, Sargano AB, Habib Z, Melanoma classifcation from dermoscopy images using ensemble of convolutional neural networks. Mathematics vol. 10, no 1, pp 2-6, 2022.

[34] Rotemberg V, Kurtansky N, Betz-Stablein B, Caffery L, Chousakos E, Codella N, et al. A Patient-Centric Dataset of Images and Metadata for Identifying Melanomas Using Clinical Context. Sci Data vol. 8, pp 81-8, 2021. doi: 10.1038/ s41597-021-00815-z.

[35] Simin AT, Baygi SM, Noori A, Cancer diagnosis based on combination of artifcial neural networks and reinforcement learning. In: 2020 6th Iranian conference on signal processing and intelligent systems (ICSPIS). IEEE, pp.1-4, 2020.

[36] Tembhurne JV, Hebbar N, Patil HY, Diwan T, Skin cancer detection using ensemble of machine learning and deep learning techniques. Multimed Tools Appl vol.16, pp.1-24, 2023.

[37] Thapar P, Rakhra M, Cazzato G, Hossain MS, A novel hybrid deep learning approach for skin lesion segmentation and classifcation. J Healthcare Eng 1709842, 2022.