

## Computationally Efficient Explainable AI Framework for Skin Cancer Detection

Kohinur Parvin<sup>1</sup>, Eshat Ahmad Shuvo<sup>\*2</sup>, Wali Ashraf Khan<sup>3</sup>, Sakibul Alam Adib<sup>3</sup>, Tahmina Akter Eiti<sup>4</sup>, Mohammad Shovon<sup>1</sup>, Shoeb Akter Nafiz<sup>5</sup>

<sup>1</sup>Department of Computer Science and Engineering, Netrokona University, Netrokona-2400, Bangladesh

<sup>2</sup>Department of Computer Science and Engineering, Bangladesh University, Dhaka-1207, Bangladesh

<sup>3</sup>Department of Computer Science and Engineering, Jahangirnagar University, Savar, Dhaka-1342, Bangladesh

<sup>4</sup>Department of Computer Science and Engineering, Siddheswari College, Dhaka-1217, Bangladesh

<sup>5</sup>Department of Computer Science & Engineering, Daffodil International University, Dhaka-1216, Bangladesh

Email(s): [kohinur.parvin@neu.ac.bd](mailto:kohinur.parvin@neu.ac.bd) (K. Parvin), [waliashrafrad@gmail.com](mailto:waliashrafrad@gmail.com) (W. A. Khan), [alamsakibul233@gmail.com](mailto:alamsakibul233@gmail.com) (S. A. Adib), [tahminaakterete13@gmail.com](mailto:tahminaakterete13@gmail.com) (T. A. Eiti), [mdshovon@neu.ac.bd](mailto:mdshovon@neu.ac.bd) (M. Shovon), [nafiz22205101898@diu.edu.bd](mailto:nafiz22205101898@diu.edu.bd) (S. A. Nafiz)

\*Corresponding Author: Eshat Ahmad Shuvo, Department of Computer Science and Engineering, Bangladesh University, Dhaka-1207, Bangladesh, [mdshuvoslk012@gmail.com](mailto:mdshuvoslk012@gmail.com)

---

### ARTICLE INFO

Article history:

Received: 17 October, 2025

Revised: 18 December, 2025

Accepted: 20 December, 2025

Online: 10 January, 2026

Keywords:

Skin Cancer, Convolutional Neural Networks (CNN), Bacterial Foraging Optimization (BFO), Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Explainable Artificial Intelligence (XAI), Grad-CAM++, LIME

---

### ABSTRACT

Skin cancer stands among some of the fastest growing and fatal malignancies of the world as a result early and accurate diagnosis of skin cancer is essential in order to enhance patient survival and treatment prognosis. Conventional methods of diagnosis including dermoscopy and histopathological examinations are expensive and time consuming also subject to inter-observer error. To address these shortcomings, this research suggests an optimized hybrid cascading framework that combines deep and machine learning strategies towards building an AI model for the detection of skin cancer. This study applied five different Convolutional Neural Network (CNN) as well as seven Machine Learning (ML) model and three optimization technique. Among them the suggested MobileNetV2 + LDA + LR model had the best accuracy of 99.33%, precision of 99.47%, recall of 97.33%, and F1-score of 98.31%. Moreover, it was found that the framework had better computational efficiency, with a time complexity of  $35.1 \pm 1.24$  seconds and moderate memory usage. Later on, Explainable AI (XAI) including Grad-CAM++ and LIME techniques were employed to validate the interpretability of the model and revealed the clinically relevant areas of lesions that affected predictions, which increased the level of transparency and credibility.

---

## 1. Introduction

Cancer has become one of the major causes of death and disability worldwide [1], costing health systems, societies and economies an enormous burden. Recent estimates show nearly 20 million new cancer cases [2] and approximately 9.7 million cancer mortality cases worldwide in 2022 and that demographic projections indicate that the number of new cases would approach 35 million per year by 2050 if current trends continue [3]. Skin cancer is turning a devastating burden to the globe currently. Specifically, the most aggressive type of skin cancer melanoma

was estimated to result in 331,722 new cases and 58,667 deaths across the globe in the latest reports [4]; once the non-melanoma skin cancers are taken into consideration, the total number of skin cancer incidences is quite large, and it belongs to the most commonly diagnosed cancer groups [5]. As a matter of fact, it has a great setback in global economy as well. Recent studies on cancer costs worldwide indicate that the economic costs of cancer are immense and concentrated: China and the United States bear the most of the shares of world cancer costs, and it is estimated that over the next few decades, cancer will cost the world about

trillions of US dollars [5,6], in terms of productivity losses and medical expenses (one prominent statistics shows that the cancer will cost the world USD 25 trillion in the next 30 years when combined with productivity losses and medical costs) [6,7]. These statistics indicates that the urgent need to implement a cost-effective prevention, early detection and effective management strategies of cancer such as skin cancer [8].

Early and precise diagnosis of skin cancer has a significant positive impact on patients' survival rates [9]. However, the traditional clinical workflows clinical inspection, dermoscopy, biopsy and histopathology have been resource consuming and exposed to inter observer variability [10]. Also, the conventional diagnosis of skin cancer required a highly qualified and experienced oncologist without which the diagnosis it quite impossible [11]. Furthermore, accessibility to specialized dermatology services is also a limitation in remote and low resourceful areas as a result diagnosis become late which also may increase the costs of treatment and may mortality rate [12]. All of these real-world limitations have encouraged the creation of automated; image-based diagnostic to assist clinicians and increase screening accessibility.

Over the last few years, deep learning (DL) specifically Convolutional Neural Networks (CNNs) have demonstrated impressive performance on dermoscopic and histopathological images tasks regarding skin cancer detection [13]. Other recent studies have reported high classification rates and investigated the use of complementary technique including attention modules, metaheuristic optimization (e.g., Ant Colony Optimization, Particle Swarm Optimization), feature selection, and Explainable Artificial Intelligence (XAI) including Grad-CAM, LIME, SHAP and many more to enhance explainability and clinical trust [14]. In spite of these advancements, there are few gaps available which actually restrict the use of an automated system to the real-world settings including:

- (i) Robustness of validation: many studies lack k-fold (10-fold) cross-validation or other commensurately rigorous statistical validation of their models in independent cohorts.
- (ii) Computational cost and complexity: few studies quantify inference time, number of parameters, memory footprint, or time and space complexity to assess their model.
- (iii) Use of optimization strategies: only a fraction of studies use systematic metaheuristic or architecture-level optimization, and systematic comparison of optimization to alternative tun Such methodological gaps decrease the trust of the reported performance and it is hard to determine the generalizability, scalability and affordability of health systems.

Based on these gaps, this study aims to design and evaluate an optimized, explainable, and computationally-sensitive hybrid cascaded deep learning framework in the detection and classification of skin cancer. The accuracy of the classification and generalization was also enhanced with higher feature selection and optimization techniques. The framework does not require large and unbalanced annotated datasets and more complex computing resources, as found in traditional methods; instead, it is specifically implemented to operate with relatively small and unbalanced datasets and with low resource requirements, therefore, simpler to fit into the standard clinical

workflow. This study would then want to make a contribution of a consistent, functionally and explainable computer aided diagnostic (CAD) model and this model not only could assist pathologists to make time sensitive and accurate diagnosis, but also address a very important gap in the environment of reliability and resource consumption [15]. Besides this, most currently implemented deep learning systems are Inconsistent black-box models, which do not provide much insight into the process of decision-making. They do not even have to be concerned with computational efficiency or scalability, which are the major concerns of a practical clinical use [16]. Indeed, as far as we know, few or, indeed, no research studies have broken down the problems of calculating their models in an explicit and analytical way (Complexity including time and space), a consideration that can debilitatingly hamper the role of transferring controlled research settings to the practice in a hospital. Overcoming these obstacles, our framework will not only be utilized to push the limits of the existing study, but also a step beyond that toward the creation of clinically practical, reliable, and sensible AI technique in the classification of skin cancer from the images. The major key contributions necessary to fill that research gaps addressed in this paper can be summarized as follows:

1. We suggest a hybrid automated system of Deep and Machine Learning in which DL is employed in terms of its effective learning-based feature extraction and ML is employed in terms of detecting skin cancer from the images.
2. The Pre-trained CNN models are applied with the purpose of extracting hierarchical features automatically and then feature selection of optimization techniques and dimensionality reduction methods are applied on the extracted features to ensure that the result is more robust and efficient.
3. The proposed model will be not only highly diagnostic but will also perform well on small datasets with minimal computing capabilities, hence it is applicable to real-time clinical practice.
4. In fact, we didn't just calculate the complexity of our proposed model but also all the remaining model's which perform comparatively higher.
5. In comparison with black-box DL models, our framework is more interpretable and reliable in the sense that it seeks to bridge the gap between AI-based predictions and clinical workflow decision-making by a pathologist.
6. Lastly we also applied XAI method including grad-cam++ and LIME in order to prove that proposed model's is more interpretable and scalable.

The headings represent the rest of our work: Section 2: Related work which illustrates the analysis of the literature of the current research based on the skin Cancer. Section 3: Methodology, the section outlines the system architecture, diagrams and requirements related to the system architecture. Section 4: Experimental Result presents the results in different tables, graphs and figures. Section 5: Conclusion and Future.

## 2. Literature Review

Several smart methods have been created in recent years to detect skin cancer based on images including: In [17], author

applied CNN algorithm to the HAM10000 dataset which had been collected from Kaggle and it has 10015 images as well as seven classes. In the image preprocessing phase this work applied the normalization and contrast enhancement techniques. The author also applied the Explainable AI techniques like GRAD-CAM and GRAD-CAM++ for making this work more reliable and understanding. This work acquired 82% classification accuracy and 0.47% loss accuracy. In [18], the study used ISIC2018 skin cancer dataset which has four classes like benign, malignant, nonmelanocytic, and melanocytic tumors and 3533 images. It used some image preprocessing techniques like image resizing, normalization and augmentation techniques. After that this research applied CNN, ResNet50, InceptionV3 and InceptionV3 & ResNet algorithms on this dataset and found the accuracy of 83.2%, 83.7%, 85.8% and 84% accordingly. In [19], the authors proposed a federated learning-based model and showed how the several hospitals exchange their private skin cancer data with a central server. The author applied DCNNs model on the three distinct dataset. They also split the dataset as a training, validation and testing dataset and for enhancing the image quality and the model's accuracy and this work also preprocessing the images like image augmentation and resizing techniques. In [20] the author applied DenseNet201 model on the Fitzpatrick 17k dataset which combines several skin condition images. This model is applied on the federated learning environment the author found 92% accuracy. The dataset was divided into 70% as training, 10% as validation and 20% as testing dataset. In [21] author applied multi class SVM model on the ISIC 2019 skin cancer dataset which has eight different classes and found the accuracy 96.25%. For preprocessing image, the author applied dull razor method for removing unwanted hair particle, gaussian filter for smoothing images and median filter for preserving edges. B. O. S. in [22] the author applied SVM and CNN model on the ISIC skin cancer dataset and found 85% accuracy. This dataset contains 23000 images and the author took 1000-1500 images for making training and testing dataset. In image preprocessing part this work removing hair, shade and glare from the images after that it did the segmentation technique. The overall summary of the literature review for this study can be summarized as follow in Table 1:

Table 1: Literature review

Ref.	Author	Method	Accuracy	Limitations
[17]	Mridha et al.	CNN	82.00%	Lower accuracy, No complexity analysis
[22]	B. O. S. Elgabbani et al.	SVM, CNN	85.00%	No Complexity, No cross validation, lower accuracy and no complexity analysis
[18]	Gouda et al.	Inception V3	85.80%	No complexity, Lower accuracy and no complexity analysis
[19]	Al-Rakhami et al.	DCNNs	90.00%	Overfitting, No cross validation and No complexity analysis
[20]	Karni et al.	DenseNet 201	92.00%	No complexity and no complexity analysis
[21]	Monika et al.	Multi class SVM	96.25%	No complexity, No cross validation and no complexity analysis

### 3. Methodology

The overall methodology of the proposed model is illustrated, in Figure 1, which is divided into five consecutive, but interdependent phases; these are the acquisition of datasets, preprocessing of the data, the architecture design of the system, model selection, and the interpretability of the proposed model with explainable artificial intelligence (XAI). The combination of these stages is to ensure that the model is more reliable, robust, and can minimize prediction errors. The data that was used in this study was collected from a secondary data source platform which contains dermatological image collections. The preprocessing stage was started with the use of advanced image enhancement methods to guarantee quality, uniform, and noise-free data. Contrast Limited Adaptive Histogram Equalization (CLAHE) was the most important of them in increasing contrast and at the same time decreasing noise in the image thus optimizing the raw dermoscopic and clinical skin photos to be further analyzed by the system [23]. Missing or inconsistent values were appropriately taken care of in order to maintain the integrity and methodological rigor of the dataset.

A set of pre-trained Convolutional Neural Network (CNN) model were applied on those preprocessed images in order to extract the features. These hierarchical based CNN models, which are characterized by their ability to extract hierarchical features, were used to break down the skin images into successively more discriminative and fine-grained feature representations, which could effectively classify the different skin cancer types. Dimensionality reduction and feature selection methods were then used to reduce redundancy, to keep only the most informative features and to reduce the computational overhead. After this step, the dataset was divided into testing (20%) and training (80%) sets thus making it easy to conduct rigorous experimentation.

Different classifiers were optimized using the training set, and this was done depending on the capability to capture complex, non-linear patterns that are considered to be linked with skin cancer, and the predictive ability of the models was confirmed with the testing set.

In order to enhance further interpretability and to bring transparency to the suggested diagnostic framework, the state-of-the-art methods of XAI, i.e., Local Interpretable Model-agnostic Explanations (LIME) and Gradient-weighted Class Activation Mapping++ (Grad-CAM++) were incorporated. LIME measures the impact of manipulations to the input information, which means that the results of the prediction are monitored in response to the alterations [24]. It then produces a linear surrogate model that reveals the most significant attributes that drive classification decisions. Instead, Grad-CAM++ generates high-resolution heatmaps, which visualize the discriminative regions of the skin lesion images, which affects the process of making a decision by the model [25]. Grad-CAM++ also provides better accuracy on tasks that have multiple target regions, compared to traditional Grad-CAM, because it computes more refined gradient-based distributions of weights over the feature maps [26]. The application of these XAI methods does not only lead to a better understanding of model behavior, but improves the clinical plausibility of the suggested skin cancer detection framework [27].

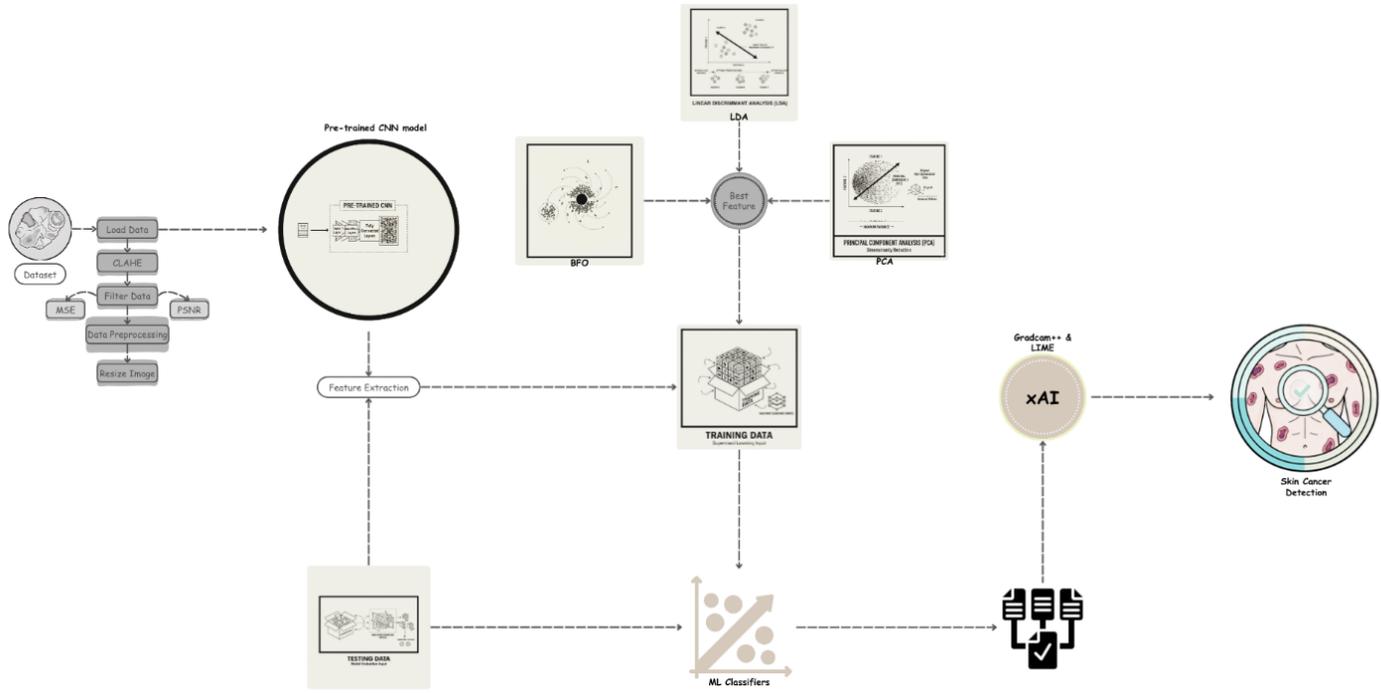


Figure 1: Workflow Diagram of this study

The framework provides a balance between predictive accuracy and explainability and, as a result, the practical interpretability requirements of dermatologists with advanced AI-driven analysis, further supporting trust and usability in the real-world healthcare environment.

### 3.1. Dataset

The dataset which were used in this study were collected from a secondary data resource platform called Kaggle [28]. A professional oncologist carefully examined and annotated the dataset and defined each image by the type and stage of lesion to guarantee reliability and clinical validity. A grading scale of 0 to 4 was used with a 0 indicating non-cancerous cases, 1 indicating Basal Cell Carcinoma, 2 indicating Malignant lesions, 3 indicating Melanoma, and 4 indicating Squamous Cell Carcinoma. The dataset is balanced between a variety of different classes, which includes 2,958 non-cancerous images, 2,093 Basal Cell Carcinoma images, 3,000 Malignant lesion images, 2,863 Melanoma images and 2,536 Squamous Cell Carcinoma images.

Table 2, provides a summary of the dataset distribution between these categories.

Table 2: Dataset Description of this study

Sample Image	Class Name	Class No.	No. of Sample image
	Non-Cancerous	0	2958
	Basal Cell Carcinoma	1	2093
	Malignant lesions	2	3000

	Melanoma	3	2863
	Squamous Cell Carcinoma	4	2536

### 3.2. Preprocessing

In order to enhance the quality of classifying skin cancer, a sequence of image processing methodologies was adopted in this research. Preprocessing is also an important part of medical image analysis because it improves the quality of data and minimizes background noise and presents models with more reliable inputs[29]. The preprocessing mechanisms used were tailor made to sharpen the skin images and facilitate extraction of their features to achieve precise classification. Some common challenges in skin images include noise, poor contrast, and unwanted artifacts that may interfere with the functionality of machine learning and deep learning models. A number of high-level preprocessing techniques were used to overcome these problems. Contrast Limited Adaptive Histogram Equalization (CLAHE) was one of the main methods, where the contrast of the images was enhanced by redistributing the intensity levels, hence emphasizing the slightest differences between cancerous and non-cancerous tissues and, by implication, are more pronounced in terms of analysis [30]. Besides contrast enhancement, noise reduction filters were also used to remove irrelevant details and enhance the meaningful tissue structures, eventually, simplifying the computation of the later models. Images that had been missed or corrupted were handled systematically to remain consistent with the data sets and image quality analysis and reconstruction methods were employed to make sure that only the data of high quality and reliability were retained. All these preprocessing measures reduced any artifacts that would affect the outcome of

the model and formed a solid base towards which machine learning and deep learning systems are cascaded to detect skin cancer, which in turn facilitated the construction of clinically viable diagnostic systems.

### 3.2.1. Image Filtering

Image filtering is a basic method of preprocessing of computer vision, which is essential in improving quality and trustworthiness of images in carrying out analysis tasks [31]. This is more crucial in the area of medical imaging where the smallest details can determine the difference between the diagnosis and the classification of the ailments. Image filtering also removes noises, highlights key features that makes images sharper, more contrasted and clearer, and may be more useful in machine learning and deep learning models [32]. The algorithm of filtering techniques which are used in this paper in order to enhance the quality of the of skin image consisted of the Gaussian filtering, median filtering and Contrast Limited Adaptive Histogram Equalization (CLAHE) etc., the random noise was removed through the use of the Gaussian filtering technique without damaging the significant edges and rendered the image smooth. Median filtering method was helpful in the fact that noise impulses were eliminated and edges in the tissue structures were preserved. Local contrast was improved through the application of CLAHE [33] to elevate the profile of minute changes in the regions of the tissue where change in intensities was minimal. By critically applying such filtering strategies, the dataset was reduced to focus on emphasis on the diagnostic patterns without distortions, which helped increase the model performance and classification rate. The initial step in converting raw skin image data to rich feature representations that can be exploited later in the medical imaging application by the image filtering. In this study there are about six types methods were used to reduce noise of the image data, including:

**Average Filtering:** Mean filtering is an image processing method that uses a sliding  $n \times n$  sized window (or kernel) of an image to remove noise and smooth out the intensity changes [34]. This kernel is usually a square (e.g.  $3 \times 3$  or  $5 \times 5$ ) and as the kernel passes through the image, the pixel at the center is replaced by an average value of the other pixels in the neighborhood. This is one of the most effective methods to reduce high-frequency noise in the image and produce a smoother one, but at the cost of blurred edges and small details. Although average filtering may help increase the average similarity of the overall image, in medical imaging, where the edges hold more importance than the overall image then this filtering technique won't be that much fruitful. The mathematical equation of average filtering is given as follow:

$$f'(x, y) = \frac{1}{n \times n} \sum_{i=-k}^k \sum_{j=-k}^k f(x+i, y+j)$$

where  $f(x, y)$  is the original pixel intensity,  $f'(x, y)$  is the filtered pixel value, and  $n \times n$  is the kernel size.

**Median Filtering:** Median filtering is another nonlinear image processing method that is extensively used in noise elimination especially in salt-and-pepper noise (also referred to as impulse noise) elimination [35]. In this method, an  $n \times n$  sliding window is moved around the image and the median of the pixel values of the surrounding is used to substitute the central pixel. In contrast

to average filtering, which generally smooths edges, median filtering still maintains sharp edges and at the same time eliminates noise, which is particularly important in medical imaging, where structural boundaries are of primary importance. The mathematical equation of median filtering can be summarized as follows:

$$f'(x, y) = \text{median} \{f(x+i, y+j) \mid (i, j) \in W\}$$

where  $W$  represents the coordinates of the window of the kernel.

**Bilateral filtering:** It is an edge preserving nonlinear and noise reducing method that uses both spatial and intensity samples [36]. Unlike the Gaussian filtering which only smooth on the basis of the proximity of the pixels, Bilateral filtering takes into consideration the similarity between the pixel intensities, thus it preserve edges, but eliminate noise. The new value of each pixel is determined as a weighted average of the surrounding pixels with the weight being dependent on the distance between the pixels and the intensity difference. This filtering technique can be mathematically described as below.

$$f'(x, y) = \frac{1}{n \times n} \sum_{i=-k}^k \sum_{j=-k}^k f(x+i, y+j) \cdot G_s(i, j) \cdot G_r(f(x+i, y+j) - f(x, y))$$

where  $G_s$  refers spatial Gaussian kernel,  $G_r$  is the range Gaussian kernel, and  $k(x, y)$  is the normalization factor.

**Min filtering:** It is also known as minimum entropy filtering as a morphological filtering method, is a process that replaces the pixel in the resulting image with the least brightness in its  $n \times n$  pixel neighborhood [37]. This technique is especially useful in eliminating bright outlier noise (like salt noise) and is most frequently used in pre-processing of medical images to eliminate undesired brightnesses at the expense of darker detail. The mathematical equation of this filtering technique is shown below:

$$f'(x, y) = \min \{f(x+i, y+j) \mid (i, j) \in W\}$$

**Max Filtering:** Unlike min filtering, max filtering substitutes a pixel in the center of the neighborhood with the max intensity of the neighborhood [37]. This method is useful in getting rid of dark outlier noise (including pepper noise) without affecting bright structures in the image. In medical imaging, max filtering may prove to be convenient in improving bright objects, e.g. highlighted tissue in a medical image, stained objects in a histopathology microscope slide. The operation is defined as:

$$f'(x, y) = \max \{f(x+i, y+j) \mid (i, j) \in W\}$$

**Gaussian Filtering:** It is a linear smoothing technique used to reduce Gaussian noise and achieve image blurring while maintaining edge information more effectively than mean filtering [38]. It employs a kernel based on the Gaussian distribution, where neighboring pixels closer to the center contribute more weight than those farther away. Even areas are smoothed and edges are preserved by this weighted averaging. The Gaussian filter is mathematically defined as:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

where  $G(x, y)$ : Gaussian kernel value at position,  $\sigma$ : standard deviation, which determines the spread of the Gaussian curve, and  $x, y$ : the position of the pixel in relation to the center of the kernel.

### 3.3. Architecture of the proposed system

Figure 1, shows the general structure of the proposed system, and the methodology is described stepwise in Algorithm 1. The pre-processing phase starts with the evaluation of image quality and removal of noise through the Contrast Limited Adaptive Histogram Equalization (CLAHE) algorithm which boosts the contrast values. Then, the images were further refined using six different filtering methods. The Image Quality Assessment Cell ( $I_{QAC}$ ) assesses the sharpness and noise level of each image according to the approach suggested by [39]. The mean square error (MSE) analysis was applied to estimate the quality predictor coefficient based on which the quality index  $Q_i$  was calculated using a linear regression model. The higher the  $Q_i$  the lower the quality of the image. Therefore, images with  $Q_{ih} = 10$  or less had to be chosen as the subject of further processing. Upon  $I_{QAC}$  validation, the validated images were transformed into sRGB space in order to normalize colors and resize them to the input demands of the AI models. The above pre-processing steps especially the quality control step will help check that only good quality images are processed in the subsequent stages [40]. Among these six different filtering techniques we choose Bilateral filtering technique for further process in our proposed model.

Following the effective preprocessing, five CNN models ResNet50, DenseNet201, InceptionV3, MobileNetV2, and Xception were deployed, and MobileNetV2 was proven to be the most effective in terms of the feature extraction capability owing to their performance and applicability to the proposed framework. To reduce redundant features before the classification process, Bacterial Foraging Optimization (BFO), Principal Component Analysis (PCA), and Linear Discriminant Analysis (LDA) were used to feature optimize. These optimized features were then introduced to seven high-performing machine learning classifiers, K-Nearest Neighbors (KNN) classifier, Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), Gaussian Naïve Bayes (GNB) and Logistic Regression (LR) classifier to classify Non-cancerous, Basal Cell Carcinoma images, Malignant lesion images, Melanoma and Squamous Cell Carcinoma.

Explainable AI (XAI) methods were added to make the predictions more interpretable and transparent. In particular, Grad-CAM++ was used to visualize class-discriminative regions in the histopathological images, showing the most significant areas that drive CNN. Additionally, by examining local pixel perturbations that affect predictions, Local Interpretable Model-Agnostic Explanations (LIME) were utilized to provide feature-level interpretability. The combination of these methods is done in such a way that the proposed framework can not only provide high-quality classification results, but also provide clinically significant visualization explanations that will increase its suitability and acceptability in medical diagnostics.

### 3.4. Algorithm

#### Algorithm 1: Algorithm of the Proposed Model

**Initialization:**  
 $P \leftarrow \{P1, P2, P3, \dots, Pn\}$   
 $B \leftarrow \{B1, B2, B3, \dots, Bn\}$   
 $C \leftarrow \{C1, C2, C3, \dots, Cn\}$

while  $I_{SRGB} \neq I_{NIL}$   
do  $I_{SRGB} = I_{QAC}(I_{RGB})$   
 $I_R \leftarrow \text{Resize}(I_{SRGB} \text{ to } 224 \times 224)$   
 $FM_i \leftarrow \text{Load feature extractor ResNet50 with parameters } P_i$   
 $f_i \leftarrow FE_i$   
 $BFO_i \leftarrow \text{Load BFO with parameters } B_i$   
 $Sf_i = BFO_i(f_i)$   
 $LDA_i \leftarrow \text{Load LDA with parameters } LDA_i$   
 $Rf_i = LDA_i(Sf_i)$   
 $PCA_i \leftarrow \text{Load PCA with defined components}$   
 $Pf_i = PCA_i(Rf_i)$   
 $CM_i \leftarrow \text{Load classifier Logistic Regression with parameters } C_i$   
 $y = CM_i(Pf_i)$

if  $y = 0$  then  
 $\psi \leftarrow \text{Non-cancerous}$   
else if  $y = 1$  then  
 $\psi \leftarrow \text{Basal Cell Carcinoma}$   
else if  $y = 2$  then  
 $\psi \leftarrow \text{Malignant Lesion}$   
else if  $y = 3$  then  
 $\psi \leftarrow \text{Melanoma}$   
else  
 $\psi \leftarrow \text{Squamous Cell Carcinoma}$   
end if  
end while  
Output:  $\psi$

Initialization:  $\alpha, \beta, \gamma, Q_{ih}$   
Load model with the  $n + 1$  parameters  $\alpha, \beta, \gamma$   
 $Q_i = \alpha + \beta \times \text{Sharpness} + \gamma \times \text{Noise}$   
Sharpness and Noise are estimated

if  $Q_i < Q_{ih}$  then  
return NIL  
end if

if  $C_{linear} \leq 0.0031$  then  
 $I_{SRGB} = 12.92 \times C_{linear}$   
else  
 $I_{SRGB} = 1.0552 \times C_{linear}^{1/2.4}$   
end if  
return  $I_{SRGB}$

Function  $BFO(F_i)$   
Initialization:  $a \leftarrow 2A - 1$ , where  $A = 1, 2, 3, 4, \dots$   
 $P \leftarrow \text{Input Image}$   
 $R_u \leftarrow \text{Respective Feature Vector}$   
 $\text{Best-Filter} \leftarrow \text{apply-Filter}(P)$   
 $Q_a \leftarrow \text{applyFilterToImage}(P, \text{BestFilter})$   
 $A \leftarrow \text{image To Array}(PQ_a)$   
for each  $A$  do  
Use( $P, Q_a$ ) to get  $R_a \mid R_a \{V_0, V_1, \dots, V_{14}\}$   
end for  
 $R_u \leftarrow R_a$   
return  $R_u$

Load PCA model with defined number of components  
 $Pf_i \leftarrow \text{Apply PCA to } Rf_i$   
return  $Pf_i$

Load LDA model with parameters  $LDA_i$   
 $Rf_i \leftarrow \text{Apply LDA to } Sf_i$   
return  $Rf_i$

3.5. Performance Evaluation Matrix

To evaluate the model’s performance four of the most dominating performance evaluation matrix were implemented those are:

**Accuracy** gives a general evaluation of how well the model is able to classify instances correctly in all classes, a broad perspective of how well the model is able to predict.

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

**Precision** is the fraction of cases of correct positive prediction among the cases which are predicted to be positive. It is the measure of how the model can reduce the false positives and mathematically determined as:

$$Precision = \frac{TP}{TP+FP}$$

**Recall (or Sensitivity)** is a standard to measure the fraction of the overall number of actual positive cases that is correctly predicted. It measures the efficiency of the model in detecting all the relevant samples and it is represented as:

$$Recall = \frac{TP}{TP+FN}$$

**F1-Score** This is the harmonic mean of Precision and Recall, which offers a balanced measure of evaluation; it is often useful when there is a class imbalance. It is calculated as:

$$F1\text{-Score} = 2 * \frac{Precision + Recall}{Precision * Recall}$$

4. Result Analysis

4.1. Comparative Analysis of Filtering Methods in Different Image Classes.

A comparative analysis of diverse image filtering methods on five types of skin lesion images, which are Non-Cancerous, Basal Cell Carcinoma, Malignant Lesions, Melanoma, and Squamous Cell Carcinoma, is presented in the Table 3. The rows are presented with a sample picture of the respective class and continue to show a sequence of filtered results produced by six filters Average, Gaussian, Median, Bilateral, Min and Max with their quantitative measures of Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR). The findings reveal that the Bilateral Filter is the best one with the lowest MSE and maximum PSNR on all the classes, which is an indication of the quality of the filter in removing noise as well as maintaining important structural information. Indicatively, the Bilateral Filter yielded MSE and PSNR values of 5.58 and 40.66 respectively in Non-Cancerous picture, 35.84 and 32.58 in Basal Cell Carcinoma picture, 11.83 and 37.40 in Malignant Lesions picture, 7.18 and 39.56 in Melanoma picture and 20.00 and 35.11 in Squamous Cell Carcinoma picture. All in all, these results prove that the Bilateral Filter is indeed the best pre-processing method to improve the quality of the image before the classification.

Table 3: Filtering Method's MSE and PSNR Values for Different Image Classes

Sample Image	Filtered Images	Class	Filtering Technique	MSE	PSNR
		Non-Cancerous	Average Filter	119.15	27.36
			Gaussian Filter	6.05	40.30
			Median Filter	9.28	38.45
			<b>Bilateral Filter</b>	<b>5.58</b>	<b>40.66</b>
			Min Filter	34.59	32.74
			Max filter	39.24	32.19
		Basal Cell Carcinoma	Average Filter	115.40	27.50
			Gaussian Filter	37.33	32.40
			Median Filter	37.55	32.38
			<b>Bilateral Filter</b>	<b>35.84</b>	<b>32.58</b>
			Min Filter	72.86	29.50
			Max filter	73.40	29.47
		Malignant lesions	Average Filter	87.09	28.73
			Gaussian Filter	15.42	36.24
			Median Filter	13.34	36.87
			<b>Bilateral Filter</b>	<b>11.83</b>	<b>37.40</b>
			Min Filter	65.38	29.97
			Max filter	62.46	30.17
		Melanoma	Average Filter	108.56	27.77
			Gaussian Filter	9.12	38.52
			Median Filter	11.63	37.47
			<b>Bilateral Filter</b>	<b>7.18</b>	<b>39.56</b>
			Min Filter	77.80	29.22
			Max filter	79.05	29.15
		Squamous Cell Carcinoma	Average Filter	93.34	28.42
			Gaussian Filter	23.55	34.40
			Median Filter	25.94	33.98
			<b>Bilateral Filter</b>	<b>20.00</b>	<b>35.11</b>
			Min Filter	78.30	29.19
			Max filter	80.06	29.09

4.2. MACS, FLOPS, Feature, and parameter values for various CNN architectures

This Table 4, compares five deep learning models, DenseNet201, MobileNetV2, ResNet50, InceptionV3 and Xception, in terms of extracted features, number of parameters, Multiply- Accumulate Operations (MACS) and Floating-Point Operations per Second (FLOPS). The Number of Features column reflects the dimensionality of the data prior to and after optimization algorithms of Bacterial Foraging Optimization (BFO) and Principal Component Analysis (PCA).

MobileNetV2 is the lightweight and resource-efficient model since it has the lowest number of parameters of 3.505 million, 327.487 million MACS, and 654.973 million FLOPS, which is much lower than the rest of the models. ResNet50, InceptionV3, and XceptionV3, by contrast, are more complex with high computational complexity of all beyond 4 billion MACS and 8 billion FLOPS, but DenseNet201 is also moderately complex with 4.390G MACS and 8.781G FLOPS. Unlike dimensionality reduction, feature optimization generates reduced dimensions with high performance and low computation cost. In general, the findings highlight that MobileNetV2 is the most appropriate model to use in the development of an efficient feature extraction and real-time skin cancer image classification because of the small number of parameters and less computational intensity.

Table 4: The features, parameter, MACS and FLOPS of various model implemented in this study

Model	Number of Features			Parameter	MACS	FLOPS
	Without Opt	BFO	PCA			
DaseNet201	1024	498	600	20.014M	4.390G	8.781G
MobileNetV2	1024	724	600	3.505M	327.487M	654.973M
ResNet50	2048	1099	800	25.557M	4.134G	8.267G
InceptionV3	2048	887	800	23.835M	5.749G	11.498G
Xception	2048	1134	800	22.855M	4.146G	8.292G

4.3. The outcome of the Skin cancer analysis without optimization

The Table 5, represents the comparative performance of different pre-trained deep learning models with traditional machine learning classifiers on skin disease classification, without optimization, in terms of Accuracy, Precision, Recall, and F1-Score measures. MobileNetV2 showed the best overall performance in terms of accuracy, with the best results of 89.71, precision of 89.25, recall of 89.73, and F1-score of 89.80, and KNN classifier with a close result 84.56. This means that MobileNetV2 offers highly discriminative features that are appropriate to work with classification. Comparatively, DenseNet201 and InceptionV3 performed fairly well on all classifiers, whereas ResNet50 kept the accuracy rates steady but rather low. The Xception model had the least performance, especially in conjunction with the Gaussian Naïve Bayes (GNB) which had the lowest accuracy of 60.37. In general, MobileNetV2 with LR or KNN worked better than other combinations, which is why it is better suited to feature extraction and classification accuracy in the context of this unoptimized configuration.

4.4. Findings from Skin Cancer Analysis With BFO Optimization

The Table 6, shows the classification of various pre-trained models with the combination of conventional machine learning classifiers with the use of Bacterial Foraging Optimization (BFO). The findings indicate that the use of BPO optimization greatly enhanced the model accuracy in most settings compared to the non-optimized one. MobileNetV2, once again, outperformed all the others with overall performance of 92.45% accuracy, 92.54% precision, 92.56% recall, and 92.84% F1-score, which shows that BFO-tuned features have strong discriminative power. KNN under MobileNetV2 was also a competitive classifier with an accuracy of 85.55%. In the meantime, ResNet50 and DenseNet201 performed steadily, but relative moderate, and InceptionV3 demonstrated a weak increase, whereas Xception was quite low in comparison with the result of the non-optimized variant. Altogether, the combination of BFO optimization significantly improved the classification performance, in particular with MobileNetV2 + LR, which proved that BFO is an effective tool that reduces the quality of features and promotes the predictive potential of hybrid deep learning models.

4.5. Findings from Skin Cancer Analysis with PCA Optimization

In the Table 7, the comparison of performance between various pre-trained CNN models together with the standard machine learning classifiers after optimization with Principal Component Analysis (PCA) is presented and measured in terms of Accuracy, Precision, Recall, and F1-Score (percentage). The findings suggest that PCA has taken a significant role in enhancing consistency of classification and discrimination of features across models. The MobileNetV2 configuration displayed the most overall results, with the Logistic Regression (LR) classifier being the highest performing with an accuracy, precision, recall, and F1-score of 91.66, 91.64, 91.66, and 91.64, respectively, and the XGBoost (XGB) classifier coming in second with 87.33% accuracy. These results prove that MobileNetV2, optimized with PCA, removes very informative and concise features that boost the performance of the classifiers. In the meantime, ResNet50 turned out to be also competitive in its performance, with SVC reaching an accuracy of 77.52% and DenseNet201 and InceptionV3 showing moderate gains over the non-optimized counterparts. Conversely, Xception was once again the worst performing classifier. In general, the use of PCA optimization was successful in minimizing the data dimensionality and redundancy, enhancing computing efficiency and classification accuracy - most prominently with MobileNetV2 + LR combination.

4.6. Findings from Skin Cancer Analysis with LDA Optimization

The Table 9, shows the classification performance of various pre-trained CNN models coupled with standard classifiers following Linear Discriminant Analysis (LDA) optimization. Optimization using LDA resulted in a visible increase in the model performance in the majority of the combinations, with MobileNetV2 again performing better than other architectures. In particular the highest results were registered by the Logistic Regression (LR) classifier that demonstrated the accuracy of 99.33, precision of 99.47, recall of 97.33, and F1-score of 98.31, which depicts a highly discriminative and generalized

representation of features. The XGBoost (XGB) and KNN classifiers of MobileNetV2 performed also with high accuracies of 88.00% and 87.64, respectively, which confirms the strength of this pre-trained model when optimized using LDA. DenseNet201 and InceptionV3, in turn, showed moderate performance gains, and ResNet50 showed relatively constant results. The Xception model, though, was still presenting the poorest results, which is indicative of its low adaptability in this regard. On the whole, the findings show that LDA optimization is an effective predictor of maximizing the separability of the classes, and the MobileNetV2 + LR version had higher predictive accuracy and generalization than any other optimized and non-optimized options. (approximately 77.7%) with higher execution time and diverse memory utilization. ResNet50+SVC model was best but still with a moderate performance (89.07% accuracy) and had the highest peak memory (3139.07 MiB). In general, MobileNetV2 + LDA + LR was the best trade-off of all model configurations in terms of classification accuracy, computational time and memory efficiency.

Table 5: The Outcome of the Skin Cancer Analysis Without Optimization

Pre-trained model	Classifier	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
DenseNet201	SVC	70.06	70.73	70.71	70.76
	RF	76.20	76.17	76.10	76.12
	DT	72.07	72.00	72.76	71.77
	GNB	72.77	73.43	72.62	72.45
	XGB	77.25	77.27	77.14	77.17
	KNN	70.76	70.74	70.67	70.65
	LR	71.57	71.47	71.47	71.47
InceptionV3	SVC	77.03	77.77	77.13	77.75
	RF	74.75	74.77	75.06	74.71
	DT	65.06	65.55	65.07	65.23
	GNB	65.17	65.13	65.35	64.74
	XGB	77.74	77.77	77.03	77.71
	KNN	70.47	70.30	70.57	70.36
	LR	76.37	76.37	76.45	76.40
MobileNetV2	SVC	83.12	82.71	83.21	82.73
	RF	80.03	80.02	87.77	87.72
	DT	82.13	82.03	82.04	82.03
	GNB	81.17	81.75	84.27	81.27
	XGB	82.75	82.54	82.77	72.64
	KNN	84.56	84.54	84.45	74.47
	LR	89.71	89.25	89.73	89.80
ResNet50	SVC	89.07	89.42	89.77	89.17
	RF	75.77	76.60	75.50	75.77
	DT	77.70	70.10	70.07	70.07
	GNB	72.77	73.67	75.37	73.65
	XGB	77.43	77.71	77.33	77.51
	KNN	77.51	77.77	77.35	77.56
	LR	77.37	77.53	77.34	77.43
Xception	SVC	77.50	77.57	77.77	77.41
	RF	76.75	76.73	77.22	76.75
	DT	70.10	70.41	70.27	70.34
	GNB	60.37	70.57	71.17	67.51
	XGB	70.62	70.62	70.72	70.57
	KNN	77.75	77.71	77.73	77.72
	LR	70.70	70.71	70.74	70.76

Table 6: The Outcome of the Skin Cancer Analysis with BFO Optimization

Pre-trained model	Classifier	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
DenseNet201	SVC	75.71	75.77	75.75	75.76
	RF	72.51	72.43	72.42	72.42
	DT	75.13	74.74	74.76	74.73
	GNB	57.70	60.41	57.75	57.67
	XGB	74.40	74.34	74.27	74.30
	KNN	73.64	73.56	73.57	73.56
	LR	75.70	75.72	75.70	75.74
InceptionV3	SVC	76.03	76.01	76.07	76.03
	RF	77.47	77.46	77.54	77.47
	DT	65.71	66.13	65.71	65.77
	GNB	56.77	57.23	57.06	57.11
	XGB	74.57	74.53	74.61	74.56
	KNN	72.12	72.01	72.27	72.07
	LR	75.20	75.70	75.12	75.32
MobileNetV2	SVC	87.20	85.21	85.21	85.23
	RF	81.33	82.33	81.37	82.32
	DT	84.13	84.13	84.14	84.13
	GNB	84.12	84.12	84.14	84.17
	XGB	87.33	87.33	87.37	87.37
	KNN	85.55	84.46	84.76	74.52
	LR	91.66	91.64	91.66	91.64
ResNet50	SVC	77.52	77.61	77.64	77.62
	RF	70.41	73.26	77.47	70.27
	DT	74.10	74.46	74.25	74.32
	GNB	44.73	47.37	50.56	45.17
	XGB	75.36	75.73	75.27	75.47
	KNN	77.25	77.47	77.14	77.31
	LR	77.07	77.10	77.21	77.16
Xception	SVC	77.57	77.60	77.77	77.54
	RF	73.47	73.35	73.76	73.30
	DT	67.72	70.37	70.03	70.17
	GNB	47.31	53.77	46.71	46.13
	XGB	76.77	76.77	77.07	76.56
	KNN	77.50	77.47	77.73	77.57
	LR	74.67	74.74	74.77	74.77

#### 4.7. Performance Evaluation of Pretrained CNN-Based Hybrid Model

It provides the comparison of the performance of different pretrained CNN-based hybrid models with their computational complexities. The results of the proposed MobileNetV2 + LDA + LR model outperformed other evaluated models with the highest accuracy of 99.33, precision of 99.47, recall of 97.33 and F1-score of 98.31, with a moderate time complexity of  $35.1 \pm 1.24$  seconds and memory consumption of PM: 2548.77 MiB, INC: 1160.05 MiB. Comparatively, other architectures like ResNet50 + SVC and DenseNet201 + BFO + LR achieved lower accuracies of 89.07% and 77.75% respectively and had a high computational requirement. On the same note, InceptionV3 + LDA + XGB and Xception + BFO + XGB showed similar results with about 77 percent accuracy, only that they took more time and consumed more resources to execute. These results show that the hybrid model implemented using MobileNetV2 provides the best accuracy, efficiency, and cost of computation, and it can be used in real-time medical image classification in resource-restricted tasks.

Table 7: The Outcome of the Skin Cancer Analysis with PCA Optimization

Pre-trained model	Classifier	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
DenseNet201	SVC	77.31	77.43	77.33	77.27
	RF	76.12	76.12	76.02	76.03
	DT	70.70	70.57	70.53	70.53
	GNB	72.30	73.67	72.54	72.27
	XGB	77.47	77.46	77.45	77.45
	KNN	76.70	76.73	76.71	76.70
	LR	77.75	77.77	77.74	77.73
InceptionV3	SVC	75.45	75.30	75.57	75.33
	RF	75.20	75.27	75.26	75.26
	DT	65.71	66.37	65.72	65.77
	GNB	54.07	54.22	54.22	53.72
	XGB	77.45	77.37	77.52	77.41
	KNN	67.47	67.22	67.62	67.23
	LR	74.37	74.40	74.44	74.42
MobileNetV2	SVC	85.12	85.71	85.21	85.73
	RF	81.03	82.03	81.77	82.72
	DT	82.13	82.03	82.04	82.03
	GNB	81.12	81.72	84.64	81.67
	XGB	85.75	82.54	82.77	72.64
	KNN	85.55	84.46	84.76	74.52
	LR	92.45	92.54	92.56	92.84
ResNet50	SVC	76.72	76.72	76.77	76.77
	RF	75.54	75.71	75.37	75.62
	DT	77.53	77.54	77.72	77.72
	GNB	70.47	71.06	73.77	71.43
	XGB	77.71	77.16	77.77	77.76
	KNN	77.11	77.17	77.11	77.14
	LR	77.64	77.72	77.47	77.67
Xception	SVC	77.34	77.33	77.61	77.27
	RF	77.17	77.16	77.47	77.14
	DT	70.71	70.76	71.13	71.02
	GNB	60.37	70.72	71.37	67.37
	XGB	77.77	77.76	70.07	77.70
	KNN	76.70	76.65	76.76	76.67
	LR	70.21	70.23	70.36	70.27

Table 8: Comparison of pretrained CNN-based hybrid models and their computational complexities

Technique	A (%)	P (%)	R (%)	F1 (%)	Time Complexity	Space Complexity
Dasenet201 +BFO+LR	77.75	77.77	77.74	77.73	59.9 s ± 1.37 s	PM: 1840.90 MiB, INC: 989.61 MiB
Inception3 +LDA+ XGB	77.75	77.56	77.56	77.55	3.22 min ± 4.11 s	PM: 2337.03 MiB, INC: 1221.06 MiB
Xception +BFO+ XGB	77.77	77.76	70.07	77.70	8.33min ± 4.2ms	PM: 2165.42 MiB, INC: 997.95 MiB
ResNet50 + SVC	89.07	89.42	89.77	89.17	2.56 min ± 2.3s	PM: 3139.07 MiB, INC 1348.95 MiB
<b>MobileNet V2 + LDA+LR</b>	<b>99.33</b>	<b>99.47</b>	<b>97.33</b>	<b>98.31</b>	<b>35.1 s ± 1.24 s</b>	<b>PM: 2548.77 MiB, INC: 1160.05 MiB</b>

Table 9: The Outcome of the Skin Cancer Analysis with LDA Optimization

Pre-trained model	Classifier	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
DenseNet201	SVC	76.66	76.62	76.61	76.61
	RF	74.73	74.73	74.65	74.67
	DT	70.67	70.45	70.35	70.36
	GNB	72.65	72.77	72.50	72.35
	XGB	76.72	76.70	76.67	76.72
	KNN	74.45	74.37	74.27	74.27
	LR	76.77	76.77	76.76	76.71
InceptionV3	SVC	77.07	76.77	76.77	76.76
	RF	73.47	73.42	73.32	73.33
	DT	74.00	74.67	73.76	74.07
	GNB	64.77	64.63	64.54	64.33
	XGB	77.75	77.56	77.56	77.55
	KNN	74.50	74.04	74.12	74.06
	LR	75.77	76.07	75.77	75.70
MobileNetV2	SVC	87.20	87.21	87.21	87.23
	RF	82.33	82.33	82.37	82.32
	DT	85.13	85.13	85.14	85.13
	GNB	85.21	85.21	85.24	85.21
	XGB	88.00	87.01	87.01	87.02
	KNN	87.64	87.65	87.66	87.65
	LR	99.33	99.47	97.33	98.31
ResNet50	SVC	77.77	77.73	77.70	77.71
	RF	75.57	75.77	75.37	75.61
	DT	70.07	70.11	70.32	70.20
	GNB	70.45	71.77	74.25	71.47
	XGB	77.74	77.74	77.77	77.76
	KNN	77.34	77.65	77.11	77.37
	LR	77.57	77.54	77.67	77.60
Xception	SVC	76.74	76.75	77.10	76.73
	RF	77.55	77.51	77.71	77.57
	DT	70.04	70.47	70.25	70.35
	GNB	67.06	67.77	67.74	66.37
	XGB	70.32	70.30	70.53	70.36
	KNN	70.71	70.77	71.25	70.75
	LR	77.53	77.71	77.66	77.67

4.8. K(10) fold cross-validation result of the proposed model

Table 10, shows the outcomes of 10-fold cross-validation conducted on the MobileNetV2 + LDA + LR model to assess its consistency and generalization performance with various data partitions. The findings indicate that, the model has very stable and excellent results across all folds with Accuracy of between 99.31 and 99.36, Precision of between 99.42 and 99.47, Recall of between 97.26 and 97.35 and F1-Score of between 98.28 and 98.34. The average standard deviation values of  $99.33 \pm 0.02\%$  (Accuracy),  $99.45 \pm 0.02\%$  (Precision),  $97.31 \pm 0.03\%$  (Recall) and  $98.31 \pm 0.02\%$  (F1-Score) demonstrate the lack of differences in the average results of the folds and illustrate the strength and stability of the model. These results prove that the suggested MobileNetV2 + LDA + LR model has a high predictive stability, good generalization and stable classification results using different data splits.

Table 10: K(10) Fold Cross-Validation Performance of the MobileNetV2 + LDA + LR Model

Fold	Accuracy (%)	Precision (%)	Recall (%)	F1-Score(%)
1	99.34	99.45	97.28	98.31
2	99.35	99.46	97.31	98.33
3	99.32	99.43	97.29	98.30
4	99.31	99.42	97.26	98.28
5	99.36	99.47	97.33	98.32
6	99.33	99.45	97.34	98.31
7	99.34	99.46	97.32	98.30
8	99.32	99.44	97.30	98.29
9	99.35	99.47	97.35	98.34
10	99.33	99.45	97.31	98.32
<b>Average ± SD</b>	<b>99.33 ± 0.02</b>	<b>99.45 ± 0.02</b>	<b>97.31 ± 0.03</b>	<b>98.31 ± 0.02</b>

4.9. Learning curve of the proposed model

As depicted in Figure 2, the learning curve of proposed MobileNetV2+LDA+LR model shows the correlation between the size of both training and cross-validation sets and the model accuracy during both training and cross-validation processes. The training accuracy first grows quickly and levels to 0.90, which means that the model learns well on the training data available. On the other hand, the accuracy of the cross-validation begins with a lower value but exhibits a consistent increasing pattern with the size of the training set, which finally approaches the value of around 0.80. The decreasing difference between the training and validation curves indicates both less overfitting and better generalization with the increase of data. The gray area around the cross-validation curve is a measure of variance between folds, which is also small at larger training sizes, again giving evidence of model stability and stability in performance.

4.10. Receiver Operating Characteristic (ROC) curve of the proposed model

Figure 3, shows the Receiver Operating Characteristic (ROC) curve of the proposed model MobileNetV2+LDA+LR, which represents the trade-off between the True Positive Rate (TPR) and False Positive Rate (FPR) of all five classes. Class 0, class 1, and class 2 ROC curves attain an area under the curve (AUC) of 1.00, which means those curves are perfect classifiers with zero misclassifications. Class 3 and class 4 also have almost perfect performances with AUCs of 0.99 and 0.94 respectively. The general findings have shown a very discriminatory model that can be successfully used to differentiate the classes. The high position of all ROC curves at the top-left corner of the graph indicates the high sensitivity and specificity, which confirms the effectiveness and stability of the proposed method in multi-class classification processes.

4.11. Confusion Matrix of the Proposed model

The confusion matrix of the proposed MobileNetV2+LDA+LR model is shown in Figure 4, and this shows the classification performance of the model in five different types of skin lesion, Non-Cancerous, B.C.C (Basal Cell Carcinoma), Melanoma, S.C.C (Squamous Cell Carcinoma), and

Malignant. The matrix demonstrates that the model perceives most of the classes with almost perfect predictions with all the diagonal values being near 1 which means that there is a high agreement between actual and predicted labels. In particular, the model accurately classifies Non-Cancerous, B.C.C and Melanoma samples in a 100 percent accuracy. In the case of S.C.C, a misclassification rate of minimal 0.0044 is present and the Malignant cases are more accurate by 0.87 with a slight confusion of 0.13 having been misclassified as S.C.C. The general distribution of values of the diagonal indicates the very good discriminative capacity and a great accuracy of the model in differentiating similar lesion types, which shows that it is a very strong and reliable model in multiclass multimodal medical image classification.

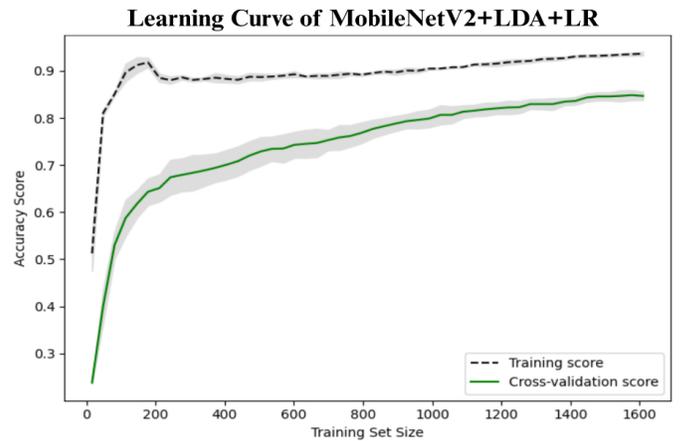


Figure 2: The proposed model's learning curve

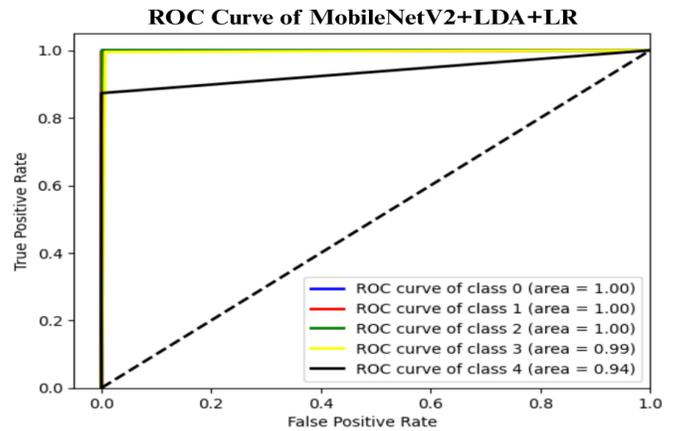


Figure 3: The proposed model's ROC curve

4.12. Outcome Explanation With XAI

Table 11 demonstrates the outcome interpretation of the proposed model using Explainable Artificial Intelligence (XAI) methods: Grad-CAM++ and LIME, with regard to four representative types of skin lesions, namely: Basal Cell Carcinoma, Malignant Lesions, Melanoma, and Squamous Cell Carcinoma. The Grad-CAM++ heatmaps are useful in that they are able to point out the discriminative areas in each lesion image with the red and yellow activation areas indicating the most important areas that the model ultimately relied on to make the final prediction. Equally, the LIME visualizations highlight which

super pixels are critical locally influencing the classification decision, and provide a more fine-grained view of the reasoning process of a model. These high correlations between the two visualization outcomes show that the model continuously targets the areas of the diagnostically significant lesions instead of the nonessential background sources. This level of interpretability does not only support the robustness and reliability of the model but also reinforces its possible application in real-world dermatological diagnostic systems by increasing clinical transparency and confidence of the decision.

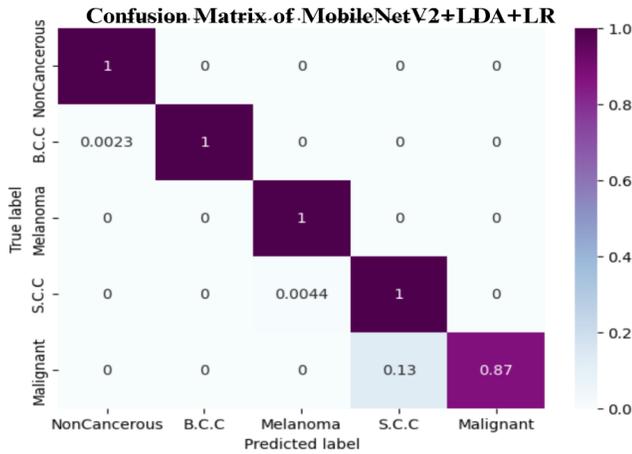
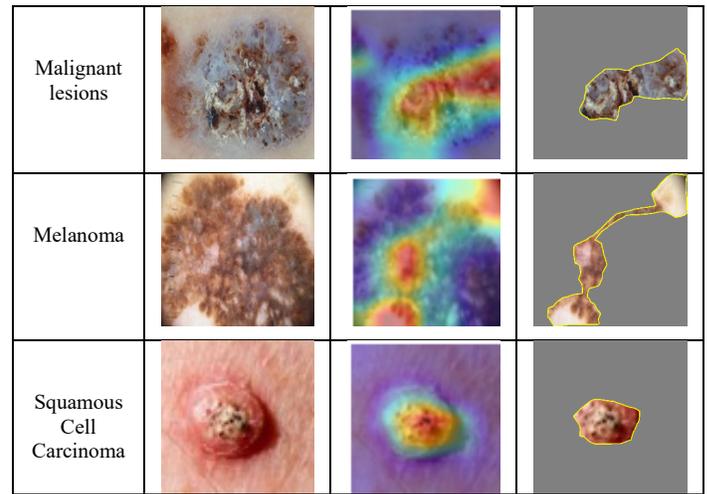


Figure 4: Confusion Matrix of the Proposed Model

Table 11: Explanation of the affected lung regions using XAI visualization

Sample Source	Actual Image	Grad-CAM++	LIME
Basal Cell Carcinoma			



#### 4.13. Discussion

A detailed comparison of the proposed MobileNetV2 + LDA + Logistic Regression model with a number of other available methods in skin cancer classification by their accuracy, optimization, explainable AI (XAI) integration, and computational complexity is provided in Table 12.

Current models, such as SVM, CNN, InceptionV3, DCNNs, and DenseNet201 reached the accuracy of between 85.00 and 96.25 with little or no optimization and XAI method. However, the suggested model has a better performance of 99.33 with the integration of various optimization models which include Bacterial Foraging Optimization (BMO), Principal Component Analysis (PCA), and Linear Discriminant Analysis (LDA). Moreover, the model uses XAI methods, namely LIME and Grad-CAM++, to improve interpretability and present visual explanations of the decision-making process. Although deep learning models are more complex, the presented approach is computationally efficient, with an execution time of  $35.1s \pm 1.24s$ , memory use of PM: 2548.77 MiB and INC: 1160.05 MiB, thus, providing a self-sustaining, reliable, and interpretable diagnostic framework in comparison to the current state-of-the-art approaches.

Table 12: Comparison between the Proposed Model and Existing Models

Ref.	Author	Method	A (%)	Opti.	XAI	Complexity	
						Time	Space
[22]	B. O. S. Elgabbani et. al.	SVM, CNN	85.00%	X	X	X	X
[18]	Gouda et al.	InceptionV3	85.80%	X	X	X	X
[19]	Al-Rakhami et al.	DCNNs	90.00%	✓	✓	X	X
[20]	Karni et al.	DenseNet201	92.00%	✓	✓	X	X
[21]	Monika et al.	SVM	96.25%	✓	✓	X	X
<b>Proposed Method</b>		<b>MobileNetV2+ LDA+ Logistic Regression</b>	<b>99.33%</b>	<b>BFO, PCA, LDA</b>	<b>LIME, Grad-CAM++</b>	<b>35.1s ± 1.24s</b>	<b>PM: 2548.77 MiB, INC: 1160.05 MiB</b>

## 5. Conclusion & Future Work

The paper suggested a streamlined hybrid deep and machine learning architecture that can achieve reliable and explainable skin cancer detection, combining MobileNetV2, Linear Discriminant Analysis (LDA), and Logistic Regression (LR). The proposed model was shown to outperform in several assessment metrics and datasets, with the best results of 99.33% accuracy, 99.47% precision, 97.33% recall, and 98.31% F1-score and with a moderate computational cost. The framework demonstrated an incredible generalization, stability and robustness to all data partitions after extensive experimentation with 10-fold cross-validation. Its efficiency was further proven by the computational complexity analysis, which also exhibited moderate time and memory usage, which made it a viable diagnostic solution, which can be used in real-time scenarios in a clinical setting, particularly in resource-limited settings. Besides providing a high performance in classification, the integration of the Explainable Artificial Intelligence (XAI) models, including Grad-CAM++ and LIME, allowed the model to give a transparent and visual representation of how it makes its decisions. These visualizations verified the fact that the model always paid attention to the clinically relevant regions of lesions, which reinforced the credibility of the model and its clinical relevance. In contrast to traditional black-box CNN models, this aspect of interpretability ensures the disconnect between AI-based analysis and clinical decision-making, increasing the level of trust in physicians and the possibility of its integration into dermatological diagnostic processes. The main contributions of the study are its balanced design that incorporated accuracy, computational efficiency, and interpretability. The study was able to overcome constraints that were experienced in previous models especially in terms of model complexity, overfitting and scalability using lightweight architectures like MobileNetV2 and dimensionality reduction techniques like LDA and PCA. Additionally, metaheuristic optimization (BFO) integration presented a powerful approach to the enhancement of feature quality and the increase in classification performance. Nevertheless, even with such a good performance, this study recognizes some limitations. The model has been trained and tested with secondary datasets, as demonstrated, which, however diverse, might not be representative of the heterogeneity of real-world clinical images. Moreover, the paper was mostly oriented on image-based data; the inclusion of multi-modal data, including patient demographics or histopathological metadata, could be added to the work to enhance the quality of the diagnoses. Future research ought to then be based on applying this framework to multi-institutional datasets, applying federated learning techniques to improve data privacy and generalizability, and how real-time can be deployed on embedded or mobile medical devices to perform point-of-care screening. Altogether, the suggested MobileNetV2 + LDA + LR framework can be regarded as an important step in the development of skin cancer diagnostics as it provides a high level of accuracy, transparency, and computational efficiency. The model does not only expand the existing body of knowledge in computer-aided dermatological diagnostics but also opens a clinically viable route towards the use of artificial intelligence in the early detection of skin cancer, finally leading to more affordable, trustworthy, and interpretable AI-based healthcare systems.

## Conflict of Interest

The authors declare no conflict of interest.

## Acknowledgment

This research received no external funding.

## References

- [1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA: A Cancer Journal for Clinicians*, **71**(3), 209–249, 2021, doi:10.3322/caac.21660.
- [2] Global Cancer Facts & Figures., <https://www.cancer.org/research/cancer-facts-statistics/global-cancer-facts-and-figures.html>, 1–5, 2025.
- [3] Cancer, Dec. 2025. [://www.cancer.org/research/cancer-facts-statistics/global-cancer-facts-and-figures.html](https://www.cancer.org/research/cancer-facts-statistics/global-cancer-facts-and-figures.html), 1–5, 2025.
- [4] F. Bray, M. Laversanne, H. Sung, J. Ferlay, R.L. Siegel, I. Soerjomataram, A. Jemal, "Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, **74**(3), 229–263, 2024, doi:10.3322/CAAC.21834.
- [5] L. Rahib, M.R. Wehner, L.M. Matrisian, K.T. Neale, "Estimated Projection of US Cancer Incidence and Death to 2040," *JAMA Network Open*, **4**(4), e214708–e214708, 2021, doi:10.1001/JAMANETWORKOPEN.2021.4708.
- [6] M. Wang, X. Gao, L. Zhang, "Recent global patterns in skin cancer incidence, mortality, and prevalence," *Chinese Medical Journal*, **138**(2), 185–192, 2025, doi:10.1097/CM9.00000000000003416.
- [7] IARC Publications Website - Skin Tumours, Dec. 2025.
- [8] S. Chen, Z. Cao, K. Prettnner, M. Kuhn, J. Yang, L. Jiao, Z. Wang, W. Li, P. Geldsetzer, T. Bärnighausen, D.E. Bloom, C. Wang, "Estimates and Projections of the Global Economic Cost of 29 Cancers in 204 Countries and Territories From 2020 to 2050," *JAMA Oncology*, **9**(4), 465–472, 2023, doi:10.1001/JAMAONCOL.2022.7826.
- [9] G. Lopes, "The Global Economic Cost of Cancer - Estimating It Is Just the First Step!," *JAMA Oncology*, **9**(4), 461–462, 2023, doi:10.1001/jamaoncol.2022.7133.
- [10] S. Chen, Z. Cao, K. Prettnner, M. Kuhn, J. Yang, L. Jiao, Z. Wang, W. Li, P. Geldsetzer, T. Bärnighausen, D.E. Bloom, C. Wang, "Estimates and Projections of the Global Economic Cost of 29 Cancers in 204 Countries and Territories From 2020 to 2050," *JAMA Oncology*, **9**(4), 465–472, 2023, doi:10.1001/JAMAONCOL.2022.7826.
- [11] S.C. Sodergren, O. Husson, S. Janssen, G.E. Rohde, M.J. Hossain, H. Abaza, A. Alkan, A. Al-Omari, I. Ben-Aharon, L. Bentsen, M.G. Guren, G. Ioannidis, H. Ishiki, M. Koehler, E. Lidington, I. Ługowska, F. McDonald, M.N. Krishnamurthy, C. Korenblum, A. Majorana, N. Memos, M. Otth, H. Pappot, M. Pérez-Campdepadrós, D. Petranovic, D. Richter, J. Roganovic, K. Scheinemann, A. Sikora-Koperska, et al., "Development of a Health-Related Quality of Life Tool for Adolescents and Young Adults With Cancer," *JAMA Network Open*, **8**(12), e2549071, 2025, doi:10.1001/JAMANETWORKOPEN.2025.49071.
- [12] E.A. Shuvo, W. Rahman, M. Shovon, P. Shaha, A. Mondol, S.H. Nahin, "Bio-inspired Heuristic Optimization-Based Cascaded Hybrid Network for Brain Cancer Screening," 2025 International Conference on Electrical, Computer and Communication Engineering, ECCE 2025, 2025, doi:10.1109/ECCE64574.2025.11013886.
- [13] T.J. Brinker, A. Hekler, J.S. Utikal, N. Grabe, D. Schadendorf, J. Klode, C. Berking, T. Steeb, A.H. Enk, C. Von Kalle, "Skin cancer classification using convolutional neural networks: Systematic review," *Journal of Medical Internet Research*, **20**(10), e11936, 2018, doi:10.2196/11936.

- [14] E.A. Shuvo, M. Shovon, M.N. Sarkar, M.S. Hossen, P. Shaha, W. Rahman, "Optimized Hybrid Cascaded Approach for Accurate Oral Cancer Detection in Histopathology Images Using Deep CNNs," 2025 2nd International Conference on Next-Generation Computing, IoT and Machine Learning, NCIM 2025, 2025, doi:10.1109/NCIM65934.2025.11160086.
- [15] Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI, Dec. 2025.
- [16] D.D. Dunipin, "Limited Access to Dermatology Specialty Care: Barriers and Tele dermatology," *Dermatology Practical & Conceptual*, **13**(1), e2023031, 2023, doi:10.5826/DPC.1301A31.
- [17] A. Esteva, B. Kuprel, R.A. Novoa, J. Ko, S.M. Swetter, H.M. Blau, S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature* 2017 542:7639, **542**(7639), 115–118, 2017, doi:10.1038/nature21056.
- [18] K. Doi, "Computer-aided diagnosis in medical imaging: Historical review, current status and future potential," *Computerized Medical Imaging and Graphics*, **31**(4–5), 198–211, 2007, doi:10.1016/J.COMP MEDIMAG.2007.02.002.
- [19] W. Samek, G. Montavon, S. Lapuschkin, C.J. Anders, K.R. Müller, "Explaining Deep Neural Networks and Beyond: A Review of Methods and Applications," *Proceedings of the IEEE*, **109**(3), 247–278, 2021, doi:10.1109/JPROC.2021.3060483.
- [20] K. Mridha, M.M. Uddin, J. Shin, S. Khadka, M.F. Mridha, "An Interpretable Skin Cancer Classification Using Optimized Convolutional Neural Network for a Smart Healthcare System," *IEEE Access*, **11**, 41003–41018, 2023, doi:10.1109/ACCESS.2023.3269694.
- [21] W. Gouda, N.U. Sama, G. Al-Waakid, M. Humayun, N.Z. Jhanjhi, "Detection of Skin Cancer Based on Skin Lesion Images Using Deep Learning," *Healthcare* 2022, **10**(7), 1183, 2022, doi:10.3390/HEALTHCARE10071183.
- [22] M.S. Al-Rakhani, S.A. AlQahtani, A. Alawwad, "Effective Skin Cancer Diagnosis Through Federated Learning and Deep Convolutional Neural Networks," *Applied Artificial Intelligence*, **38**(1), 2024, doi:10.1080/08839514.2024.2364145
- [23] A. Karni, Q. Abbas, J. Ahmad, A.K.J. Saudagar, "Skin cancer classification using novel fairness based federated learning algorithm," *PeerJ Computer Science*, **11**, e3171, 2025, doi:10.7717/PEERJ-CS.3171/TABLE-5.
- [24] M. Krishna Monika, N. Arun Vignesh, C. Usha Kumari, M.N.V.S.S. Kumar, E. Laxmi Lydia, "Skin cancer detection and classification using machine learning," *Materials Today: Proceedings*, **33**, 4266–4270, 2020, doi:10.1016/J.MATPR.2020.07.366.
- [25] B. ELGABBANI, M.F.- ARTIFICIAL, undefined 2025, "ARTIFICIAL INTELLIGENCE BASED DERMAL CANCER DETECTION USING COMPUTER VISION," *Hunandaxuexuebao.ComBOS ELGABBANI, MA FARAJARTIFICIAL INTELLIGENCE*, Dec. 2025, doi:10.5281/ZENODO.14909049.
- [26] D. Mane, O. Khode, S. Koli, K. Bhat, P. Korade, "CNN-Based Medical Image Restoration Using Customized Adaptive Histogram Equalization," *Lecture Notes in Electrical Engineering*, **1098**, 267–285, 2024, doi:10.1007/978-981-99-7383-5\_21.
- [27] M.T. Ribeiro, S. Singh, C. Guestrin, "Why should i trust you?" Explaining the predictions of any classifier," *DI.Acm.OrgMT Ribeiro, S Singh, C GuestrinProceedings of the 22nd ACM SIGKDD International Conference on Knowledge*, 1135–1144, 2016, doi:10.1145/2939672.2939778.
- [28] Skin Cancer Images, Dec. 2025.
- [29] G. Brancaccio, A. Balato, J. Malvey, S. Puig, G. Argenziano, H. Kittler, "Artificial intelligence in skin cancer diagnosis: a reality check," *ElsevierG Brancaccio, A Balato, J Malvey, S Puig, G Argenziano, H KittlerJournal of Investigative Dermatology*, **144**(3), 492–499, 2024, doi:10.1016/J.JID.2023.10.004.
- [30] A. Chattopadhyay, A. Sarkar, P. Howlader, V.N. Balasubramanian, "Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks," *Ieeexplore.Iee.OrgA Chattopadhyay, A Sarkar, P Howlader, VN Balasubramanian2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 839–847, 2018, doi:10.1109/WACV.2018.00097.
- [31] E.D. Pisano, S. Zong, B.M. Hemminger, M. DeLuca, R.E. Johnston, K. Muller, M.P. Braeuning, S.M. Pizer, "Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms," *SpringerED Pisano, S Zong, BM Hemminger, M DeLuca, RE Johnston, K Muller, MP BraeuningJournal of Digital Imaging*, 1998•Springer, **11**(4), 193–200, 1998, doi:10.1007/BF03178082.
- [32] M.S. Hossen, A. Shuvo, A. Arif, P. Shaha, A. Rahman, M. Saiduzzaman, F. Al Farid, H.A. Karim, A. Saleh, M. Miah, "An Efficient Deep Learning Framework for Brain Stroke Diagnosis Using Computed Tomography (CT) Images," *arXiv preprint arXiv:2507.03558*, 2025.
- [33] E.A. Shuvo, W. Rahman, M.S. Hossain, M.T. Islam, M.S. Iqbal, "Bio-inspired Heuristic Optimization-based Cascaded Network for Diabetic Retinopathy Screening," *Ieeexplore.Iee.OrgEA Shuvo, W Rahman, MS Hossain, MT Islam, MS Iqbal2024 3rd International Conference on Advancement in Electrical and*, 2024, 2024, doi:10.1109/ICAEEEE62219.2024.10561821.
- [34] Digital Picture Processing - Azriel Rosenfeld, Avinash C. Kak - Google Books, Dec. 2025.
- [35] D. Lai, V. Verfaillie, J.P.-O. Engineering, undefined 1985, "Median filter as image preprocessor for machine recognition," *Spiedigitalibrary.OrgDC Lai, V Verfaillie, J PotenzaOptical Engineering*, **24**(6), 1985, doi:10.1117/12.7973624.
- [36] K. He, J. Sun, X. Tang, "Guided image filtering," *Ieeexplore.Iee.OrgK He, J Sun, X TangIEEE Transactions on Pattern Analysis and Machine Intelligence*, **35**(6), 1397–1409, 2013, doi:10.1109/TPAMI.2012.213.
- [37] S.D. Deshpande, M.H. Er, R. Venkateswarlu, P. Chan, "Max-mean and max-median filters for detection of small targets," *Spiedigitalibrary.OrgSD Deshpande, MH Er, R Venkateswarlu, P ChanSignal and Data Processing of Small Targets*, 3809, 74–83, 1999, doi:10.1117/12.364049.
- [38] D. Marr, E.H.-P. of the R.S. of, undefined 1980, "Theory of edge detection," *Royalsocietypublishing.OrgD Marr, E HildrethProceedings of the Royal Society of London. Series B*, 207(1167), 187–217, 1980, doi:10.1098/RSPB.1980.0020.
- [39] I.S. Maksymov, I. Staude, A.E. Miroshnichenko, Y.S. Kivshar, "Optical yagi-uda nanoantennas," *Degruyterbrill.Com*, **1**(1), 65–81, 2012, doi:10.1515/NANOPH-2012-0005/HTML.
- [40] M. Sonka, V. Hlavac, R. Boyle, "Image pre-processing," *Image Processing, Analysis and Machine Vision*, 56–111, 1993, doi:10.1007/978-1-4899-3216-7\_4.

**Copyright:** This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).