Journal of Carcinogenesis
Covering Genesis to Therapeutics

Journal of Carcinogenesis, ISSN: 1477-3163
2025; Vol 24: Issue 9s, https://carcinogenesis.com/

Early Detection of Skin Cancer Using Machine Learning Approach

Zulkifli Muhammad Kalgo[1], Ahmad Baita Garko[1,2], Muhammad Saidu Aliero[3]

[1]Department of Computer Science, Kebbi State University of Science and Technology Aliero, Nigeria

[2]Department of Computer Science, Federal University Dutse, Nigeria

[3]Department of Computer Science, Kebbi State University of Science and Technology Aliero, Nigeria

ABSTRACT

Over the past decade, a skin cancer detection has attracted a lot of research attention in the both medical and soft computing domains because of the increasing number of applications for early detection of skin diseases. Strong evidence suggests that skin cancer is one of the top three deadly cancers brought on by Deoxyribo Nucleic Acid (DNA) damage is skin, which can be fatal. Cells begin to grow uncontrollably because of this damaged DNA, and they are currently expanding quickly. There have been several studies done on the automated detection of cancer in photographs of skin lesions. The existing studies failed to properly address the problem of cancer detection accuracy. Therefore, this study aimed to propose a framework that combined data from various inputs to improve the accuracy of skin cancer detection. The proposed approach uses feature importance mapping and mining to convert a multidata input into a single output classification model, which has advantages over statistical approaches. Five distinct machine learning algorithms and model prediction results were used to evaluate the recommended solution. The results showed that the suggested method can collect data, analyze it, and make predictions with high accuracy, ranging from 85 to 99.6%.

Keywords: *Skin cancer detection, machine learning skin cancer, computerized skin disease detection*

How to Cite: Zulkifli Muhammad Kalgo, Ahmad Baita Garko, Muhammad Saidu Aliero, (2025) Early Detection of Skin Cancer Using Machine Learning Approach, *Journal of Carcinogenesis*, *Vol.24, No.9s*, 627-637 .

1. INTRODUCTION

One of the deadliest types of skin cancer, malignant melanoma, is responsible for 75% of all cancer-related fatalities[1]. Skin cancer is a serious health issue. Over the past several decades, melanoma incidences have rapidly increased in Europe, America, and Australia [2]. Both melanoma and non-melanoma are becoming more common; among persons over 50, the yearly incidence of melanoma is increasing by 0.6%, while the anticipated number of new cases of cutaneous melanoma in 2016 was 76,380, or 4.5% of all new cancer cases [3]. According to projections, there will be about 96,480 new instances of melanoma identified in the US alone, 7,230 deaths from the disease, and that number is likely to increase to 150,000 [4]. Early diagnosis, however, clearly depends on patient attentiveness and a physician's precise assessment. There is little information on the test procedures, and the diagnostic variances are wide enough . Inspection is the main foundation of the diagnosis. Conventional imaging simply records what the human eye can see with a digital camera, but dermoscopy, also known as epiluminescence light microscopy (ELM), requires an experienced specialist to obtain the necessary picture. Researchers have created computational algorithms for automated diagnosis that rely on quantitative metrics in an effort to increase diagnosis reliability [5]. These technologies include machine learning, which can assist medical professionals identify the afflicted areas more quickly and objectively by facilitating quantitative judgment in addition to their expertise. Particularly, Deep learning helps to perform predictive analysis that offers a wide range of information to assess a particular problem and make decision in automated way with high accuracy. Deep learning is a form of modern technology developed by researchers such as deep learning convolutional Mobilenet [5]. Mobilenet has been shown to perform better than expert dermatologists at skin cancer detection. Biological processes such as that the nerve cells present in the brain and how they interact with each other encouraged the development of a look-alike artificial neural network technology called the CNN. A CNN is skilled at learning fast from images that it “gets”, learns the patterns, colours and continues to teach itself from what it has learned to improve its performance (a process known as machine learning) (Wibowo et al., 2020). Machine learning utilizes different algorithms among which neural networks has become a very popular tool in the software industry [6]. The existence of neural networks can easily be seen in many digital services like the recommended system.

With regard to image-based applications, the recent growth in strength and popularity of deep learning can also be portrayed in applications on skin cancer, which relies on a multi-stage processing of images to extract increasingly more complex features from them [6, 7]. As the features get extracted at higher levels, they tend to progressively represent more distinctive and integrated characteristics of the image, thus resulting in high performances for image classification [8].

However, the existing standard dermoscopic datasets contain images of light-skinned people, mostly from Europe, Australia, and the United States. For accurate skin cancer detection in dark- skinned people, a neural network must learn to account for skin color [9]. However, doing so is possible only if the neural network observes enough images of dark-skinned people during the process of training. Therefore, datasets having sufficient lesion images of dark-skinned and light-skinned people are necessary for increasing the accuracy of skin cancer detection systems.

2. LITERATURE REVIEW

Deep learning is an intuitive process whose complexity of learning increases with the increase in the number of layers. Due to its high performance, it is regarded as a mature application for medical diagnostics. In recent times, deep learning has contributed significantly for skin lesion classification problems [10]. However, limited data set creates tougher environment for the potential groundbreaking research in medical diagnostics with deep learning. One reason is dependency of the deep learning algorithm on training data size as it requires millions of parameters and large amount of labeled data to learn [11]. When limited data is used to train deep learning model, it uses large amount of its resources to train the model, creating over setting issues.

Skin cancer detection is significantly aided by deep neural networks. They are made up of a collection of connected nodes. Their interconnected neural structure is comparable to that of the human brain. Their nodes collaborate to find solutions to specific issues [12]. After being taught for certain tasks, neural networks become authorities in the fields in which they were trained. Earlier research in this area used deep learning and machine learning to identify skin conditions. Using hierarchical structures, [13] studied the use of deep learning for skin cancer diagnosis. They found that skin lesions had a hierarchical organization that dermatologists take into account when diagnosing them. Automatic methods, however, do not utilize this data, executing the diagnostic in a one-vs.-all manner, where all forms of lesions are taken into consideration. [14] has present a new approach for the task of hair removal on dermoscopic images based on deep learning techniques. The proposed model relies on an encoder-decoder architecture, with convolutional neural networks, for the detection and posterior restoration of hair's pixels from the images. [15] uses a Dermoscopic photos from the ISIC archive to construct a dataset with two classes, which was then used to categorize benign and malignant cancer kinds. The goal is to raise the categorization score because of early diagnosis. To achieve this, dermoscopic images obtained from a dataset are subjected to several image preparation processes, such as color clarity, edge detection, and noise extraction. Following this processing operation, these processed photos are classified using the InceptionV2 deep learning network. This study's findings show that pre-processing boosts accuracy ratio by 3.33 points, yielding an accuracy rate of 88.66%. while [16] uses a variety of deep convolutional neural network models that have already been trained, including resnet18, squeezenet, google net, vgg16, and vgg19, and applies them to two separate datasets. MODE-NODE and ISIC skin lesion datasets are the datasets. According to testing accuracy measurement, the study concludes that vgg19 is the most suitable DCNN and reached 98.8%. The research conducted by [17] suggested employing CNN for the detection of skin cancer, In their study, they take into account this technique for raising melanoma detection precision. However, there aren't any sizable datasets of optical lesion images that can be used to train deep networks for dermatoscopic modality. In order to get over these constraints, they propose a transfer learning model to train a collection of deep learning networks and a novel CNN architecture. The convolutional neural network, which produces a detection rate of 97%, achieves the best detection rates. However, improvement was required in the area of detection accuracy, in another research conducted by [18] suggested a lesion indexing network (LIN) based on DL. By extracting more features, they used DL-based LIN to achieve good results. To further improve results, segmentation performance had to be improved.

The author of [19] employed CNN to identify pigmented melanocytic lesions as skin cancer from dermoscopic pictures. Screening for non-melanocytic and non-pigmented skin cancer was challenging, nevertheless. Additionally, its detection accuracy decreased. In other research by [20] proposed a DCNN that includes three phases that work incredibly well to identify skin lesions. To improve contrast, color transformation is initially applied; lesions Lesion boundaries are then extracted using the CNN technique, and deep features are extracted using transfer learning. However, the strategy produced good results in some instances of dataset, but outcomes may vary in other instances. In [21], the author developed a CNN-based model to identify melanoma skin cancer. To enhancing the image prior to and following segmentation, respectively, pretreatment and postprocessing of the image were utilized. By fusing local and global contextual data, the model generated lesion regions. It achieved a decent classification and prediction performance.

The current approach for cancer skin detection using facial detection and machine learning is bias to dark skinned people when deployed in dark skinned environment, which make the approach not to be reliable when tested on dark skinned people, to improve the accuracy of detection, an extensive dark-skinned dataset needs to be included in model training to archive higher performance during cross validation. Considering model or framework that can predict skin cancer in multi racial environment can save life, resources and establish common understanding in the research community. If further fill the gap of current trend more especially in African countries.

3. PROPOSED FRAMEWORK

Figure 1 provides a broad overview of the suggested framework, which comprises the training data, new instances as input parameters. The fact that skin cancer detection with artificial intelligence is not as effective as it could be is one of the biggest issues. Before the system can effectively evaluate and interpret the features from picture data, it must go through a rigorous training process that takes a lot of time and extremely powerful hardware.

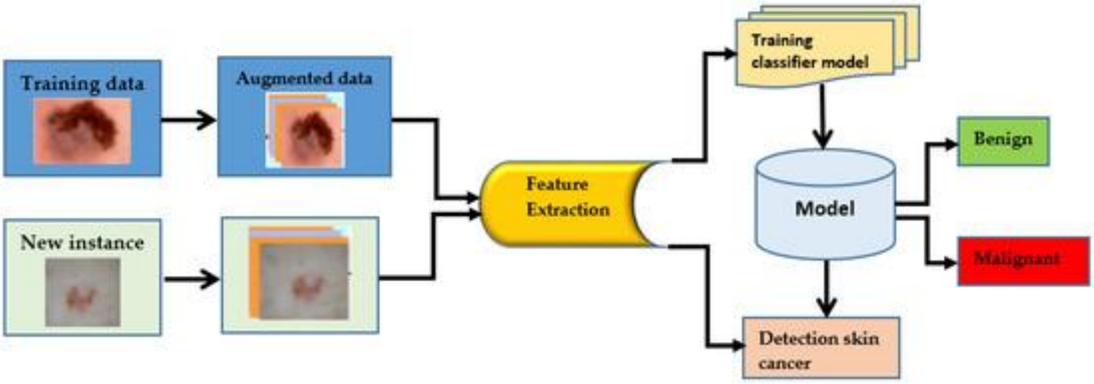


Figure 1: proposed framework

Similarly, the proposed framework implements skin cancer feature extraction to analyze dermoscopic images by extracting features like texture, color, and structure to identify malignant lesions. Then the model automatically learns these features to achieve higher accuracy in detection and classification for early diagnosis and treatment (see Figure 1).

3.1 Machine learning model

A branch of intelligent technology known as "machine learning" is concerned with creating computational models and algorithms that can acquire knowledge compared to information that's themselves. There are two main approaches to machine studying: managed learning and unsupervised study. Supervised learning uses algorithms that have been trained to execute tasks using labeled data, requiring both the data being input and the associated anticipated outputs, while unsupervised learning does not rely on labeled data and is used to carry out tasks that extract knowledge by analyzing the data despite established labels. To further clarify the performance of the model five possible algorithms have been selected for investigation. These models are popular and often serve as baseline standards, despite being less sophisticated and exciting than several of those recent innovations in this field. These models also have the benefit of being fundamental solutions for a wide range of other applications beyond skin cancer prediction, making them suitable for ML libraries. The scikit-learn Python library is used in both implementations in this study, and the library documentation contains information regarding the default method parameters [26]. A high-level summary of the selected ML models and their predicted outcomes is provided in the remaining portion of this section.

3.2 Model Training and Testing

To maximize both local and global feature representation, we develop a model for skin lesion classification model. Early on, the model uses ConvNeXtV2 blocks to effectively capture hierarchical features; later on, it switches to separable self-attention layers to capture long-range relationships. With larger kernels, depth-wise separable convolutions, and layer normalization, ConvNeXtV2 improves on traditional CNNs, allowing the model to capture subtle details and complex patterns that are crucial for differentiating benign from malignant lesions. This enhances the model's capacity to identify minute irregularities and more precisely categorize skin tumors. Later on, separable self-attention takes the place of normal self-attention, which maintains global dependency capture while decoupling the spatial and channel dimensions to reduce computing complexity. The model is quite useful for a variety of skin cancer classifications because of this approach, which concentrates on diagnostically relevant regions, reduces background noise, and improves sensitivity and specificity. This method ensures accurate classification of skin lesion types by fusing the versatility of attention mechanisms with the computational efficiency of convolutional networks.

When ML algorithms are used to generate predictions on data, datasets are usually divided according to the train and test ratio during the model training process to assess the precision of predictions. It is a simple and fast technique that aids in evaluating the results of ML algorithms and selecting the approach that best suits the model prediction challenge. The method involves dividing the original dataset into training and test ratios, such as 60:40, as shown in Figure 2. The model is matched using the first part, referred to as the training dataset. The second part, called the test dataset, is fed into the model to test predictions and assess the results of those predictions.

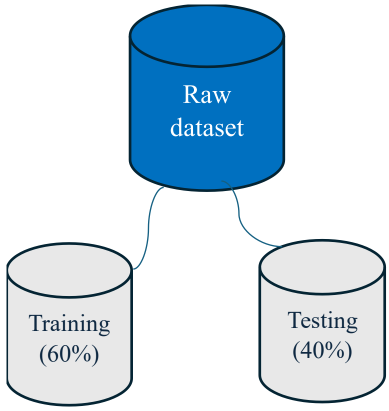


Figure 2: dataset for training and testing

3.3 Performance Evaluation Merics

Accuracy

One statistic used to assess the algorithm's prediction's % reliability is accuracy (A). The proportional ratio of many accurate guesses to all predictions is how it is established. For instance, the accuracy may be computed employing the following formula after incorporating the confusion metric.

(1)

Precision

Precision (P) is another metric measurement that tells you how often a correct forecast turns out to be accurate. It is calculated as the ratio of all correct positive predictions. One way to express a mathematical accuracy metric is as follows:

(2)

Recall

Recall (R) is like accuracy, recall tells you what percentage of the positive points were projected to be positive. According to equation 3, it is the proportion of true positive results to the sum of true positive result and false negative results.

(3)

F-Score

F-score Finally, an overall metric that measures one point value score to communicate the equilibrium between accuracy and recall is crucial for assessing our model. As a result, it makes considerably easier to combine the accuracy and recall parameters; the F-score or F-Measure is the typical method for doing so (see equation 4). (4)

(4)

4. EXPERIMENT

The study considered two potential ML models which are well-known and frequently serve as efficiency baselines, despite being less sophisticated and exciting than several of the more recent advances in this area of study. These models also have the benefit of being fundamental solutions for a wide range of other applications beyond skin cancer prediction, making them suitable for ML libraries. The scikit-learn Python library is used in both implementations in this study, and the library documentation contains information regarding the default algorithm parameters [26]. To further improve the final prediction and stabilize the learning process, a Model Exponential Moving Average (EMA) was employed. The final model was more reliable and continuously performed well thanks to this smoothing process. In order to provide a fair foundation for evaluating different models and architectures, the 224×224 pixel resolution for input photos was chosen in accordance with accepted standards in both dermatological image analysis and the larger computer vision literature.

4.1 Data collection and Ethic of Sensitive Data

The ISIC 2019 dataset, a well-respected and varied benchmark resource frequently utilized to further skin cancer detection research, is used in this study. This dataset, which covers a wide range of skin lesion types, is the perfect starting point for thoroughly assessing the efficacy and resilience of cutting-edge diagnosis moves in. Our approach is based on a state-of-the-art hybrid deep learning framework that combines separable self-attention mechanisms with ConvNeXtV2 blocks in a seamless manner. This advanced architecture provides unmatched sensitivity and specificity in detecting and categorizing skin malignancies by fusing the accuracy of local feature extraction with the capacity to recognize global contextual patterns. The model also uses ViTs to improve its ability to analyze fine details and contextual correlations in dermoscopic imagery. ViTs are enhanced by sophisticated data augmentation techniques and transfer learning processes. In order to guarantee the study's reproducibility and its value to the larger scientific community, we offer a thorough explanation of the model's development, deployment, and training procedures. It is hoped that this thorough documentation will stimulate additional developments in cancer detection techniques and spur innovation in the medical imaging industry.

Figure 3 presents a summary of the data distribution across categories, showcasing five representative photos per class from the ISIC 2019 dataset. The collection includes 25,331 dermoscopic images with labels that are categorized into Melanoma, Melanocytic Nevus (NV), BCC, Actinic Keratosis (AK), BKL, DF, Vascular Lesion (VASC), and Squamous Cell Carcinoma (SCC) are the eight different types of skin lesions. There is a lot of variances in the pictures. in resolution, which spans 101 distinct resolution sizes and ranges from 576×768 to 1024×1024 pixels. Every image uses three RGB channels and is in full color. The stark class disparity in this dataset is one of its main problems. For instance, there are around 51 times as many photos in the NV class as in the VASC class, which can distort model training and reduce classification accuracy. Advanced strategies including data augmentation, weighted loss functions, and complex sampling approaches are required to address this mismatch. This work responds by utilizing a CNN-ViT hybrid model, which combines the global pattern recognition skills of ViTs with the local feature extraction strengths of CNNs. The dataset imbalance is successfully addressed by this integrated strategy, which results in improved performance and reliable skin lesion categorization.

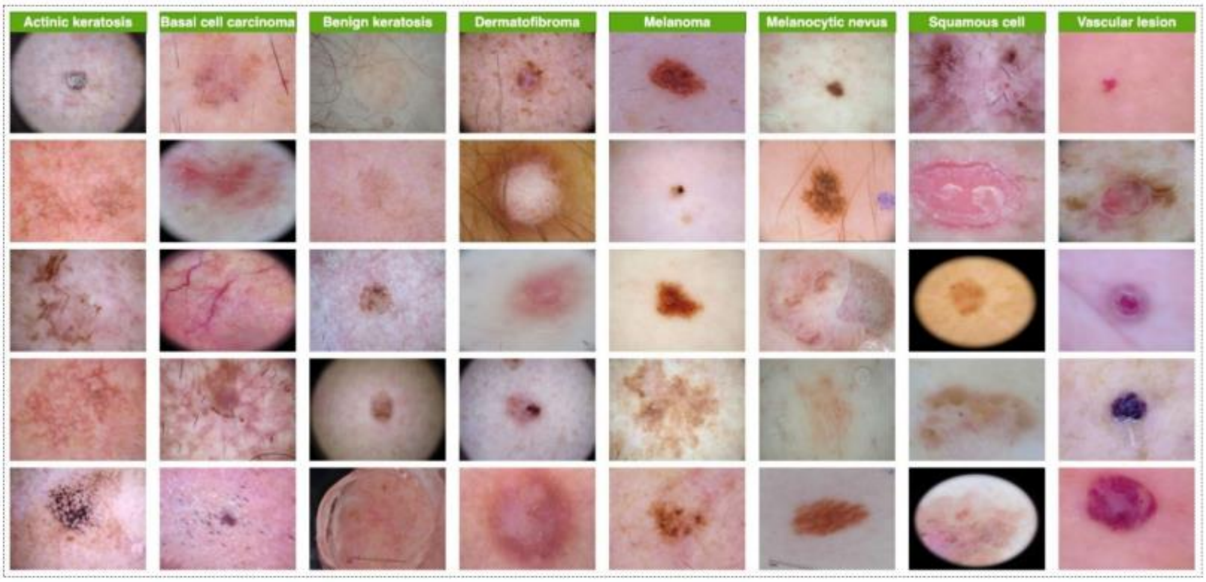


Figure 3: Representative sample images from each class in the ISIC 2019 dataset

Advanced strategies including data augmentation, weighted loss functions, and complex sampling approaches are required to address this mismatch. This work responds by utilizing a CNN-ViT hybrid model, which combines the global pattern recognition skills of ViTs with the local feature extraction strengths of CNNs. ViTs' global pattern recognition skills and improved performance and robust classification of skin lesions are the results of this integrated strategy, which successfully tackles the dataset imbalance. The dataset imbalance is successfully addressed by this integrated technique, which improves performance and provides reliable skin lesion categorization.

4.2 Random Forest

Leo Breiman and Adele Cutler created the popular ML algorithm Random Forest (RF), which aggregates the output of many decision trees to produce a single outcome. Because it can handle both classification and regression issues, its popularity has been spurred by its versatility and ease of use as well as its efficacy as a random forest classifier. This article will explain the random forest algorithm's operation, benefits, and methods for random forest regression. The Random Forest model ability to deal with datasets with continuous parameters as in regression and categorical parameters as in classification is among its most crucial characteristics. It does better on problems involving regression and classification. The proposed model prediction results is presented in Table 1 using RF algorithm. As can be seen, the model was able to reach reliable prediction with accuracy of from 85% to 97% (see Table 1).

Table1: Result of skin cancer prediction using RF model

Score Bin	Positive Rate	Negative Rate	Fraction Above Threshold	Accuracy	F1 Score	Precision	Recall	Negative Precision	Negative Recall	Cumulative AUC
(0.900,1.000)	1462	1	0.570	0.807	0.988	0.999	0.978	0.970	0.999	0.000
(0.800,0.900)	10	0	0.576	0.852	0.992	0.999	0.986	0.981	0.997	0.001
(0.700,0.800)	0	0	0.576	0.883	0.992	0.999	0.986	0.981	0.997	0.001
(0.600,0.700)	4	0	0.578	0.913	0.994	0.999	0.990	0.986	0.996	0.003
(0.500,0.600)	6	0	0.583	0.956	0.995	0.999	0.995	0.994	0.994	0.005
(0.400,0.500)	0	51	0.583	0.976	0.995	0.954	0.995	0.994	0.994	0.005
(0.300,0.400)	4	79	0.585	0.930	0.997	0.892	0.999	0.999	0.992	0.006
(0.200,0.300)	1	139	0.589	0.855	0.995	0.801	1.000	1.000	0.986	0.013
(0.100,0.200)	0	105	0.596	0.87	0.989	0.744	1.000	1.000	0.969	0.029
(0.000,0.100)	**0**	**404**	**1.000**	**0.583**	**0.736**	**0.583**	**1.000**	**1.000**	**0.000**	**0.999**

4.2 K-Nearest Neighbours

K-Nearest Neighbors (KNN) is a supervised ML technique that may be applied to regression challenges as well as classification tasks. It produces predictions based on the average value (for regression) or the majority class (for classification) after identifying the "k" data points that are the closest to a given input. KNN is a non-parametric, instance-based learning technique as it does not. The proposed model shows solid prediction results with reliability of 80% to 96% accuracy using KNN algorithm (see Table 2)

Table2: Result of skin cancer prediction using KNN model

Score Bin	Positive Rate	Negative Rate	Fraction Above Threshold	Accuracy	F1 Score	Precision	Recall	Negative Precision	Negative Recall	Cumulative AUC
(0.900,1.000)	728	1	0.570	0.807	0.988	0.999	0.978	0.970	0.999	0.000
(0.800,0.900)	84	0	0.576	0.852	0.992	0.999	0.986	0.981	0.999	0.001
(0.700,0.800)	57	0	0.576	0.883	0.992	0.999	0.986	0.981	0.999	0.001
(0.600,0.700)	57	0	0.578	0.913	0.994	0.999	0.990	0.986	0.999	0.003
(0.500,0.600)	79	0	0.583	0.956	0.995	0.999	0.995	0.994	0.999	0.005
(0.400,0.500)	70	51	0.583	0.966	0.995	0.954	0.995	0.994	0.933	0.005
(0.300,0.400)	12	79	0.585	0.930	0.997	0.892	0.999	0.999	0.832	0.006
(0.200,0.300)	0	139	0.589	0.855	0.995	0.801	1.000	1.000	0.653	0.013
(0.100,0.200)	0	105	0.596	0.799	0.989	0.744	1.000	1.000	0.519	0.029
(0.000,0.100)	**0**	**404**	**1.000**	**0.583**	**0.736**	**0.583**	**1.000**	**1.000**	**0.000**	**0.000**

4.3 Support Vector Machine

A supervised ML approach for classification and regression applications is called Support Vector Machine (SVM). It looks for the optimal hyperplane, or border, between the data's various classes. When doing binary categorization, such as spam vs. non-spam or cat vs. dog, it is helpful. Maximizing the variance across the two classes is SVM's primary objective. The more accurate the model does on fresh, untested data, the greater the margin. The SVM algorithm's main concept is to maximize the margin across two classes to determine which hyperplane best divides them. The gap between the hyperplane and the closest data points on each side is known as the margin. The proposed model prototype prediction demonstrates higher reliability of prediction when trained with SVM algorithm with accuracy of 70% to 86% (see Table 3).

Table3: Result of skin cancer prediction using SVM model

Score Bin	Positive Rate	Negative Rate	Fraction Above Threshold	Accuracy	F1 Score	Precision	Recall	Negative Precision	Negative Recall	Cumulative AUC
(0.900,1.000)	950	1	0.510	0.837	0.932	0.999	0.874	0.850	0.999	0.999
(0.800,0.900)	44	13	0.533	0.849	0.955	0.999	0.914	0.893	0.999	0.982
(0.700,0.800)	30	14	0.549	0.855	0.970	0.999	0.942	0.925	0.999	0.964
(0.600,0.700)	28	3	0.564	0.865	0.983	0.999	0.968	0.957	0.999	0.960
(0.500,0.600)	16	16	0.573	0.867	0.991	0.999	0.983	0.976	0.999	0.940
(0.400,0.500)	18	25	0.591	0.860	0.992	0.985	0.999	0.999	0.979	0.908
(0.300,0.400)	1	89	0.602	0.833	0.984	0.968	1.000	1.000	0.954	0.793
(0.200,0.300)	0	167	0.626	0.808	0.964	0.931	1.000	1.000	0.896	0.579
(0.100,0.200)	0	220	0.648	0.706	0.946	0.898	1.000	1.000	0.842	0.297
(0.000,0.100)	**0**	**231**	**1.000**	**0.583**	**0.736**	**0.583**	**1.000**	**1.000**	**0.000**	**0.941**

4.4 Artificial Neural Network

Computer programs called artificial neural networks, (ANNs), are made to simulate how the human brain interprets data. ANN employ neurons to evaluate data, find patterns, and formulate predictions, much how the brain uses neurons to process information and make choices. These networks are made up of layers of linked neurons that collaborate to find solutions to challenging issues. The main thesis is that, like our brains, ANN may "learn" from the data they process. Applications for them range from picture recognition to customized recommendation systems. In this post, we will examine more about ANNs, how they operate and other key topics. ANNs operate by using a process known as training to identify patterns in data. By contrasting its predictions with what occurs throughout training, the network modifies itself to increase its accuracy. The proposed model prediction results are presented in Table 4 using ANN algorithm. As can be seen, the model was able to reach reliable prediction with accuracy of from 80% to 96% (see Table 1).

Table4: Result of skin cancer prediction using ANN model

Score Bin	Positive Rate	Negative Rate	Fraction Above Threshold	Accuracy	F1 Score	Precision	Recall	Negative Precision	Negative Recall	Cumulative AUC
(0.900,1.000)	783	1	0.510	0.807	0.988	0.999	0.720	0.978	0.970	0.999
(0.800,0.900)	37	13	0.533	0.852	0.992	0.999	0.754	0.986	0.981	0.997
(0.700,0.800)	24	14	0.549	0.883	0.992	0.999	0.776	0.986	0.981	0.997
(0.600,0.700)	23	3	0.564	0.913	0.994	0.999	0.798	0.990	0.986	0.996

(0.500,0.600)	18	16	0.573	0.956	0.995	0.999	0.814	0.995	0.994	0.994
(0.400,0.500)	12	25	0.591	0.966	0.995	0.954	0.825	0.995	0.994	0.994
(0.300,0.400)	39	89	0.602	0.930	0.997	0.892	0.861	0.999	0.999	0.992
(0.200,0.300)	120	167	0.626	0.855	0.995	0.801	0.971	1.000	1.000	0.986
(0.100,0.200)	31	220	0.648	0.799	0.989	0.744	1.000	1.000	1.000	0.969
(0.000,0.100)	**0**	**231**	**1.000**	**0.583**	**0.736**	**0.583**	**1.000**	**1.000**	**1.000**	**0.000**

4.5 Decision Tree

A decision tree maps out several options and their potential results, which aids in decision-making. It is applied to machine learning tasks such as prediction and classification. We'll learn more about decision tree structures, their kinds, and other fundamental ideas. By displaying several possibilities and their relationships, a decision tree assists us in making decisions. The root node, which represents the complete dataset, is the first major question in its tree-like structure. Based on characteristics in the data, the tree then branches out into other options. Finally, the proposed model prediction results using DT are presented in Table 5 using RF algorithm. As can be seen, the model was able to reach reliable prediction with accuracy of from 80% to 94% (see Table 1).

Table5: Result of skin cancer prediction using DT model

Score Bin	Positive Rate	Negative Rate	Fraction Above Threshold	Accuracy	F1 Score	Precision	Recall	Negative Precision	Negative Recall	Cumulative AUC
(0.900,1.000)	783	1	1	0.807	0.837	0.932	0.999	0.807	0.988	0.999
(0.800,0.900)	37	13	13	0.852	0.849	0.955	0.999	0.852	0.992	0.999
(0.700,0.800)	24	14	14	0.883	0.855	0.970	0.999	0.883	0.992	0.999
(0.600,0.700)	23	3	3	0.913	0.865	0.983	0.999	0.913	0.994	0.999
(0.500,0.600)	18	16	16	0.931	0.867	0.991	0.999	0.956	0.995	0.999
(0.400,0.500)	12	25	25	0.943	0.860	0.992	0.985	0.966	0.995	0.954
(0.300,0.400)	39	89	89	0.930	0.833	0.984	0.968	0.930	0.997	0.892
(0.200,0.300)	120	167	167	0.855	0.808	0.964	0.931	0.855	0.995	0.801
(0.100,0.200)	31	220	220	0.799	0.706	0.946	0.898	0.799	0.989	0.744
(0.000,0.100)	**0**	**231**	**231**	**1.000**	**0.583**	**0.736**	**0.583**	**0.583**	**0.736**	**0.583**

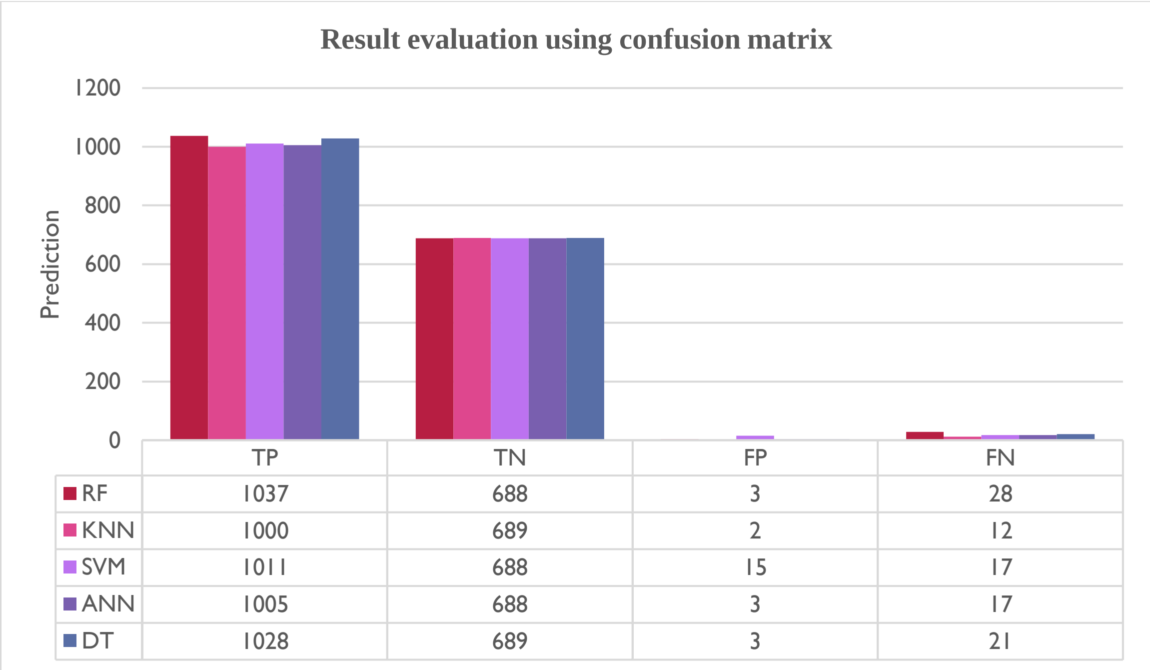
4.6 Overall Model Performance Evaluation using Confusion Matric

The primary purpose of a confusion matrix is to show how well a ML classifier predicts values from a sample dataset whose true values are unknown. This method is quite easy to comprehend. Refer to Table 6 for prediction analysis using confusion matric. It is important to note that FP and FN indicate inaccurate model prediction in the notations, while TP and TN indicate correct model prediction.

	Positive	Negative
Predicted Positive	True positive (TP)	False positive (FP)
Predicted Negative	False negative (FN)	True negative (TN)

Given some crucial ratios, the matrix (see Table 10) can be understood as below.
True Positive Result= (Total TP/Actual TP)
True Negative Result= (Total TN/Actual TN)
False Positive Result= (Total FP/Actual TP)
False Negative Result = (Total FN/Actual TN)
The visual illustration (see Figure 7) shows the five methods that were evaluated on the model using various discrimination thresholds; the TP, TN, FNR, and FP are depicted in the Figure for various threshold terms; in ML, the TP and TN are referred to as the probability of successful identification, and the FNR and FPR are referred to as the probability of an error. The confusion matrix assessment offers methods to choose potential ideal models for prediction and remove less effective separately from the expense of setting or the distribution of classes. This assessment is linked to a straightforward and typical statistical efficiency gauge in classification theory and hypothesis validation, and it is similar to the manner in that

the costs and benefits analysis of assessment selection is done.
Five techniques were used to evaluate the prediction efficiency of the trained model, as shown in Figure 7. The KNN and ANN algorithms are most impacted, even though particular algorithms exhibit minimal unexpected behavior (FP and FN). Overall, the training and prediction models operate exceptionally well. It is important to note that, in comparison to trained RF models in [16], the efficiency of RF, DT, SV, KNN and ANN models are intriguingly enhanced. The clustered training and prediction greatly lowered the FPR for all models and produced a more accurate TPR for prediction



	TP	TN	FP	FN
RF	1037	688	3	28
KNN	1000	689	2	12
SVM	1011	688	15	17
ANN	1005	688	3	17
DT	1028	689	3	21

The suggested model predicts 1037/1040 (99.7%) TP and 688/690 (99.7%) TN correctly out of real accurate predictions, but 3/690 (0.4%) FN and 28/1040 (2.6%) FP on RF are wrong predictions. For DT, 1028/1040 (98.8%) TP and 689/690 (99.8%) TN are the accurate model predictions while 3/690 (0.4%) FN and 21/1040 (2%) were the incorrect predictions. Similarly, the correct model prediction on KNN have been observed as 1000/1040 (96.1%) TP, 689/690 (99.8%) TN, and incorrect prediction of 2/690 (0.3%) FN, and 12/1040 (1.1%) on. Similarly, for ANN the correct model prediction can be observed as 1005/1040 (96.6%) TP, and 688/690 (99.7) while the incorrect prediction are 3/690 (0.4%) and 17/690 (1.6%) accordingly. Finally correct prediction for SVM are 1011/1040 (97.2%) TP, 688/690 (99.7%) TN, and incorrect prediction of the model are 15/690 (0.2%) and 17/690 (2.4%) accordingly

5. DISCUSSION AND CONCLUSION

This study addresses the important problem of early and precise skin cancer diagnosis, a major worldwide health issue where prompt detection is essential to enhancing patient survival and treatment results. There are significant categorization challenges due to the intrinsic visual similarities between benign and malignant tumors. An inventive hybrid deep learning model that combines separable self-attention mechanisms with ConvNeXtV2 blocks to improve feature extraction and classification performance is put forth in order to overcome these issues. When used in the early phases, ConvNeXtV2 blocks are made to efficiently record subtle patterns and complex local features, which are crucial for differentiating across visually identical lesion types. Later on, the inefficiencies of conventional self-attention processes are overcome by using separable self-attention to reduce computational complexity and concentrate diagnostically important areas. Using sophisticated data augmentation and transfer learning approaches, the model was rigorously trained and validated using the ISIC 2019 dataset, which included eight different skin lesion categories, to guarantee robustness and reliability. With an F1-score of 91.82%, 93.48% accuracy, 93.24% precision, and 90.70% recall, the architecture showed exceptional performance. Under controlled experimental settings, these outcomes outperform over 20 cutting-edge deep learning models, including CNN-based and ViT-based architectures. With just 21.92 million parameters, the model is computationally efficient despite its remarkable performance, which makes it ideal for real-time and mobile application deployment. The Proposed Model has substantial potential for clinical adoption and better patient outcomes by tackling the major issues of feature extraction, computational efficiency, and classification accuracy. This establishes a new benchmark for dependable and scalable skin cancer diagnosis.

Future research should aim to validate the model using datasets from diverse clinical environments, ensuring that it generalizes well across varied contexts. Domain adaptation techniques could also be explored to further enhance the model's adaptability. Although the proposed hybrid model has demonstrated exceptional accuracy and computational efficiency, there are a number of limitations that should be acknowledged in order to maximize its clinical applicability and robustness. One of the study's major limitations is its reliance on the ISIC 2019 dataset, which, despite its comprehensiveness, may not fully capture the variability present in real-world clinical settings. Resolving the dataset's class imbalance presents another difficulty. Despite the use of data augmentation and transfer learning to address this problem, misclassification rates are still greater for some categories, such BKL and melanoma, which have a significant visual similarity. Future research should look into cutting-edge methods like adaptive loss functions that give more weight to underrepresented classes or generative adversarial networks (GANs) for creating synthetic data in order to get around this.

Another noteworthy drawback is the lack of explainability features. The interpretability of AI models in clinical practice is crucial for fostering professional trust and guaranteeing moral application. It may be easier to integrate the model into diagnostic procedures and boost clinician confidence in its recommendations if explainable AI (XAI) techniques like saliency maps and attention heatmaps are used to make the model's predictions more clear. Although the suggested model is computationally efficient when compared to other cutting-edge architectures, little research has been done on how well it performs on devices with limited resources, including embedded computers or older mobile phones. In order to lower the model's computational footprint and enable wider deployment in low-resource contexts or portable diagnostic tools, future studies should concentrate on improving the model using approaches like pruning, quantization, or knowledge distillation.

The model's current emphasis on classification tasks restricts its ability to provide thorough lesion analysis. Its clinical relevance could be further increased by adding lesion segmentation or localization to its list of skills. A multitasking framework that incorporates these features might offer a comprehensive diagnostic approach. Its accuracy and adaptability could also be increased by investigating its use with additional diagnostic modalities like histopathology or sophisticated imaging techniques.

Another encouraging avenue is the model's real-time implementation in telemedicine and mobile healthcare systems. Its usefulness in enhancing healthcare outcomes and accessibility might be demonstrated by assessing its performance in real-time diagnostic settings, particularly in underprivileged or remote areas. Additional validation would be beneficial for such applications in order to guarantee dependability in practical settings. Finally, future research endeavors ought to incorporate ethical and regulatory considerations. The appropriate application of AI in healthcare requires resolving potential biases, protecting data privacy, and adhering to medical device laws. To close the gap between research and real-world application, it will be essential to establish precise standards for the clinical adoption of such technologies.

REFERENCES

1. Ali, M. and T. Aittokallio, *Machine learning and feature selection for drug response prediction in precision oncology applications.* Biophys Rev, 2019. **11**(1): p. 31-39.
2. Ang, J.C., et al., *Supervised, Unsupervised, and Semi-Supervised Feature Selection: A Review on Gene Selection.* IEEE/ACM Trans Comput Biol Bioinform, 2016. **13**(5): p. 971-989.
3. Marcos-Zambrano, L.J., et al., *Applications of Machine Learning in Human Microbiome Studies: A Review on Feature Selection, Biomarker Identification, Disease Prediction and Treatment.* Front Microbiol, 2021. **12**: p. 634511.
4. Wang, F., et al., *Navigating the microarray landscape: a comprehensive review of feature selection techniques and their applications.* Front Big Data, 2025. **8**: p. 1624507.
5. A, A., *A Deep Learning Approach to Skin Cancer Detection in Dermoscopy Images.* J Biomed Phys Eng, 2020. **10**(6): p. 801-806.
6. Ahmadi Mehr, R. and A. Ameri, *Skin Cancer Detection Based on Deep Learning.* J Biomed Phys Eng, 2022. **12**(6): p. 559-568.
7. Ahmed, B., et al., *Malignant Melanoma: Skin Cancer-Diagnosis, Prevention, and Treatment Skin Cancer Detection TechnologyUltraviolet Radiation Exposure and Its Impact on Skin Cancer Risk.* Crit Rev Eukaryot Gene Expr, 2020. **30**(4): p. 291-297.
8. Álvarez-Salafranca, M., et al., *Dermoscopy in Basal Cell Carcinoma: An Updated ReviewDevelopment of a non-melanoma skin cancer detection model Artificial Intelligence for Skin Cancer Detection: Scoping Review.* Actas Dermosifiliogr (Engl Ed), 2021. **112**(4): p. 330-338.
9. Arnold, M., et al., *Global Burden of Cutaneous Melanoma in 2020 and Projections to 2040.* JAMA Dermatol, 2022. **158**(5): p. 495-503.
10. Debelee, T.G., *Skin Lesion Classification and Detection Using Machine Learning Techniques: A Systematic Review.* Diagnostics (Basel), 2023. **13**(19).

11. Dildar, M., et al., *Skin Cancer Detection: A Review Using Deep Learning Techniques.* Int J Environ Res Public Health, 2021. **18**(10).
12. Zhu, Y., et al., *Macrophage autophagy deficiency-induced CEBPB accumulation alleviates atopic dermatitis via impairing M2 polarization SkinLiTE: Lightweight Supervised Contrastive Learning Model for Enhanced Skin Lesion Detection and Disease Typification in Dermoscopic Images A multimodal vision foundation model for clinical dermatology.* Cell Rep, 2023. **42**(11): p. 113430.
13. Zhou, L., et al., *Global, regional, and national trends in the burden of melanoma and non-melanoma skin cancer: insights from the global burden of disease study 1990-2021.* Sci Rep, 2025. **15**(1): p. 5996.
14. Zafar, M., et al., *Skin Lesion Analysis and Cancer Detection Based on Machine/Deep Learning Techniques: A Comprehensive Survey.* Life (Basel), 2023. **13**(1).
15. Y, B.V., et al., *Modified Le-Net Model with Multiple Image Features for Skin Cancer Detection A multimodal vision foundation model for clinical dermatology.* Cancer Invest, 2025. **43**(7): p. 485-514.
16. Watson, M., D.M. Holman, and M. Maguire-Eisen, *Ultraviolet Radiation Exposure and Its Impact on Skin Cancer Risk.* Semin Oncol Nurs, 2016. **32**(3): p. 241-54.
17. Schadendorf, D., et al., *Melanoma*
A novel skin cancer detection model using modified finch deep CNN classifier model. Lancet, 2018. **392**(10151): p. 971-984.
18. Rojas, K.D., et al., *Skin cancer: Primary, secondary, and tertiary prevention. Part II Skin cancer in skin of colorSkin cancer--early detection.* J Am Acad Dermatol, 2022. **87**(2): p. 271-288.
19. Rezk, E., M. Eltorki, and W. El-Dakhakhni, *Improving Skin Color Diversity in Cancer Detection: Deep Learning Approach.* JMIR Dermatol, 2022. **5**(3): p. e39143.
20. Reynaud, B., et al., *Multiple pigmented skin metastases of breast cancer mimicking cutaneous melanoma Developing skin cancer education materials for darker skin populations: crowdsourced design, message targeting, and acral lentiginous melanomaBurden of skin cancer in Colombia.* Ann Dermatol Venereol, 2022. **149**(3): p. 219-220.
21. Prelaj, A., et al., *Artificial intelligence for predictive biomarker discovery in immuno-oncology: a systematic review Improving Skin Color Diversity in Cancer Detection: Deep Learning Approach.* Ann Oncol, 2024. **35**(1): p. 29-65.