##### A Project Report

On

**Skin Caner Detection**

**Using Machine Learning**

##### Bachelor of Technology in

**Computer Science & Business System**

**Submitted b***y*

AYUSH DAS (13031122032)

MARCH 2024



**Techno Main**

**EM-4/1, Sector-V, Salt Lake**

**Kolkata- 700091**

**West Bengal**

**India**

## ABSTRACT

The discovery of skin cancer at an early stage leads to more favorable treatment results. The combination of visual inspection and dermoscopy works only through clinician-dependent observation, yet biopsy analysis provides invasive confirmation of diagnoses. Advanced technologies, which include AI machine learning together with multispectral and confocal microscopy, help achieve better accuracy levels. Computer algorithms perform image analysis while reaching more than 90% accuracy, which surpasses dermatologist capabilities. These noncontact screening instruments make massive population testing possible through smartphone applications, which prove helpful for populations with limited access to healthcare. Advanced technologies face data bias problems along with regulatory hurdles, yet demonstrate potential to alleviate skin cancer by advancing equal and timely diagnostic capabilities.

The incorporation of AI diagnostic technology with telemedicine provides remote assessment by experts through innovative diagnostic solutions. Dermatological care deficiencies become bridged by expert remote evaluations through these solutions. Cloud-based platforms allow physicians to conduct real-time image examination that provides simultaneous expert opinions while lowering incorrect biopsy rates and reducing unnecessary biopsies. The system obtains better performance by continuously learning with various data sets. The system learns to reduce bias, thus enhancing its ability to serve various skin types effectively. Users become capable of monitoring their skin health by utilizing informational public outreach and mobile apps designed for simplicity. People can use this technology for proactive monitoring of their skin health condition.

As these technologies evolve, the successful deployment of early skin cancer detection systems demands active cooperation between medical clinicians, engineers, and government decision-makers. Ethical deployment methods and the development of trust, along with early skin cancer detection, require collaboration between clinicians and engineers working with policymakers to make detection more accessible and effective worldwide.

**Keywords: Machine Learning, Image Classification, Neural Networks, Medical AI, Dermoscopic Analysis, Computer**

## CONTENTS

Declaration

Acknowledgments

Abstract

1. Introduction 8-9
   1. [Skin Cancer](#_TOC_250008) 8
   2. [Motivation](#_TOC_250007) 8
   3. [Problem Statement](#_TOC_250006)  8-9
   4. Objectives 9
2. Literature Review 10-12
3. Proposed Methodology and techniques 13-16
   1. Techniques used 13-14
      1. Logistic Regression 13
      2. K-Nearest Neighbor 13
      3. Naïve Bayes 13
      4. Decision Tree 13
      5. Random Forest 14
      6. Support Vector Machine 14
   2. [Proposed Method 1](#_TOC_250001)4-15
   3. [Materials used 1](#_TOC_250000)5-16
   4. ​

Pre-processing

3.4.1 Normalization

3.4.2 Standardization

1. **Result & Discussion** **17-22**
   1. Model Performance Evaluation 17-22
      1. Confusion Matrix Analysis 17
      2. Other Metrics 17-18
      3. Performance Analysis of Logisitc Regression 19
      4. Performance Analysis of Decision Tree 20
      5. Performance Analysis of k-Nearest Neighbor 20
      6. Performance Analysis of Naïve Bayes 21
      7. Performance Analysis of Random Forest 21-22
      8. Performance Analysis of Support Vector Machine 22
   2. Discussion 22
2. **Conclusion and Future Work 23-24**

**6 References 25-27**

# List of Figures

1.Flowchart of the proposed method 15

2.Skin images extracted from HAM10000 dataset 16

3.Performance of three matrices used 19

# List of Tables

1. Comparison of performance among other models 18

**Chapter-1**

**Introduction**

## 1.1 Skin Cancer

## 

## Melanoma represents the most feared skin cancer type worldwide because it spreads into other body regions when doctors do not identify it during its early stages [1]. Skin cancer also manifests in two less dangerous forms known as basal cell carcinoma (BCC) and squamous cell carcinoma (SCC) yet both demand prompt medical attention [2]. The early discovery of skin cancer through proper detection leads to superior therapeutic outcomes that enhance patient survival chances [3]. To identify suspicious lesions dermatologists historically depend on the traditional ABCDE criteria that includes dimensions of Asymmetry, Border irregularity, Color variation, Diameter greater than 6mm, and Evolving nature. Automated diagnosis systems remain essential because clinical detection depends on subjective assessments from medical experts [4].

## 1.2 Motivation

The worldwide occurrence of skin cancer continues to rise because of increasing incidence rates even though public awareness efforts continue to grow[5]. The early discovery of skin cancer remains essential because it leads to better survival probabilities while making treatments simpler[6]. The current diagnostic procedures depend on dermatologist visual inspections and invasive biopsy procedures yet these methods face issues with subjective evaluations and restricted access and patient discomfort problems[7].

Advanced technological tools have brought about revolutionary changes in the techniques used to detect skin cancer[8]. Deep learning-based computer vision algorithms display powerful abilities to analyze images showing patterns which human observation cannot detect in dermatology[9]. Medical studies demonstrate that artificial intelligence technology performs melanoma diagnoses with over 90% accuracy which proves superior than dermatologist expertise in controlled environments.[10]

## 1.3 Problem Statement

Current diagnostic procedures encounter multiple built-in obstacles. Skin cancer diagnosis accuracy shows wide variation depending on the diagnostic skill level of clinicians because minor visual variations between benign and malignant tissues lead to assessment challenges. The procedure of acquiring confirmatory biopsies leads to diagnostic problems that include both postoperative infections and scarring and healthcare budget strain and additional resource requirements.

There was a common method of skin self-examination for detecting skin lesions in first stage, it's the method of the indications "ABCDE". However, some researchers suggest limitations of this technique and doubt its efficiency. It was revealed that respondents could not distinguish and see the differences between benign skins, such as nevi or moles. Research on the early detection of nodular melanoma has shown that nodular melanoma skin lesions do not sometimes follow the ABCD criteria. A check of the self-diagnosed visual images of a patient's skin concluded that an inexperienced person will have difficulty applying the "ABCDE" indications without the right images.

## 1.4 Objectives

This document examines the emerging landscape of skin cancer detection technologies, evaluating their potential impact on clinical practice, patient outcomes, and healthcare systems. By exploring both the promising capabilities and implementation challenges of these innovations, we aim to provide a comprehensive overview of technology's role in addressing the global burden of skin cancer.

**Chapter 2**

**Literature Review**

In recent years, deep learning has significantly advanced the field of skin cancer detection, offering diagnostic accuracy comparable to or surpassing that of dermatologists. This literature review explores various models and approaches developed to enhance lesion classification, segmentation, and interpretability. The studies highlight innovations in CNN architectures, transfer learning, multimodal systems, and data augmentation strategies to address clinical challenges and dataset limitations.

**Machine Learning approaches**:

**Kassem et al.** [11] designed a model which applied techniques from CNN to extract features that later entered a SVM classical classifier. The strategy proved superior to independent CNN applications in particular testing conditions thereby demonstrating benefits from uniting traditional methods with deep learning approaches. (2020)​

**Bistroń and Piotrowski** [12] conducted a comparative analysis of machine learning algorithms for skin cancer diagnosis. They evaluated Logistic Regression, k-Nearest Neighbors, Naive Bayes, Decision Tree, Random Forest, and Support Vector Machine classifiers. The study found that Random Forest achieved the highest accuracy and recall among the traditional machine learning models. (2022)

**Nguyen et al.** [13] performed a comparative analysis of machine learning algorithms for melanoma diagnosis using dermoscopic images. They assessed Logistic Regression, Decision Trees, Random Forest, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and CNN. While CNN outperformed other models, traditional machine learning algorithms like Random Forest and SVM showed promising results. (2024) ​

**Deep Learning approaches**:

**Nasr-Esfahani et al.** [14] utilised the AlexNet model which underwent pre-training to detect melanoma through high accuracy outcomes despite low quantities of available data. (2016)​

**Esteva et al.** [15] employed GoogleNet Inception v3 CNN for their study which utilized ImageNet pretrained weights and processed more than 129,000 skin lesions images. Deep learning technology utilized in their model delivered results at levels equivalent to those of board-certified dermatologists for skin cancer diagnosis. (2017)​

**Yu et al.** [16] developed a deep residual network for lesion segmentation, which, when followed by classification, enhanced accuracy by focusing the model on the lesion rather than surrounding skin. (2017)​

**Zhang et al.** [17] introduced a method to generate visual explanations for CNN predictions, using attention maps to highlight regions influencing the model’s decision. Such techniques bridge the gap between complex models and clinical trust. (2018)​

**Brinker et al.** [18] carried out a systematic assessment of deep learning approaches for melanoma recognition. The researchers discovered that CNN-based approaches showed superior performance compared to typical machine learning approaches (including Support Vector Machines and decision trees) and human dermatologist diagnosis capabilities in particular cases. Standardized evaluation protocols must be established according to the study because they guarantee fair model comparison. (2019)​

**Tschandl et al.** [19] combined clinical metadata consisting of patient age, sex and lesion location information with image data through their multimodal approach. Clinical metadata combined with visual data led to better diagnosis classification according to their research results which indicated an enhanced diagnostic precision. (2019)​

**Mahbod et al.** [20] conducted research to determine how changing image resolution levels affects CNN performance outcomes. The researchers discovered that better image resolution enhanced detection accuracy however increased processing requirements which should be evaluated for actual deployment. (2020)​

**Gessert et al.** [21] combined multiple CNN architectures (e.g., ResNet, DenseNet) to achieve state-of-the-art results on the ISIC 2018 challenge dataset. This approach leverages the strengths of diverse models to improve overall performance. (2020)​

**Amin et al.** [22]proposed a deep feature fusion approach combining CNN-based feature maps for better localization and classification of skin cancer. The model significantly improved detection performance on benchmark datasets. It addressed challenges like intra-class variability and low contrast. Their hybrid design showed robustness across various lesion types. The study emphasized the importance of both texture and shape cues. (2020)​

**Bakkouri and Afdel** [23] introduced a multi-layer feature fusion network tailored for dermoscopy image analysis. The model improved feature extraction by integrating shallow and deep layers. It achieved high accuracy in differentiating benign from malignant lesions. Their method demonstrated enhanced interpretability. The system was tested on ISIC datasets with promising results. (2020)​

**Wei et al.** [24] developed an ensemble of lightweight CNNs aimed at mobile/edge deployment. Despite its compact architecture, the system maintained high classification accuracy. It used selective fusion of predictions from multiple CNNs. This design offered a balance between detection power and computational efficiency. It was ideal for real-time skin cancer screening applications. (2020)​

**Kaymak et al.** [25] performed an extensive comparison of fully convolutional networks (FCNs) for lesion segmentation. Several architectures like U-Net and FC-DenseNet were evaluated. The research highlighted the strengths and weaknesses of each model. It provided insights into lesion boundary preservation and segmentation accuracy. The findings support tailored model choice based on lesion complexity. (2020)​

**Khan et al.** [26] proposed a hybrid pipeline combining saliency-based preprocessing and optimal CNN feature selection. Their approach reduced irrelevant background noise and enhanced lesion representation. They tested various classifiers, finding CNNs to outperform traditional models. Their pipeline achieved significant gains on HAM10000 and ISIC datasets. The method also supported explainability in decision-making. (2020)​

**Jain et al.** [27] compared six pre-trained models using transfer learning for skin lesion classification. Among VGG19, InceptionV3, and others, Xception achieved the highest accuracy. The authors emphasized dataset balancing and preprocessing. They also explored model fine-tuning for performance gains. The study concluded that deeper architectures outperformed shallower ones. (2021)​

**Kausar et al.** [28] built an ensemble model by combining predictions from five deep CNNs. The fusion method leveraged strengths of each architecture, including DenseNet and ResNet. It achieved state-of-the-art accuracy on the ISIC dataset. Their method mitigated overfitting and improved generalization. They also implemented cross-validation for reliability. (2021)​

**Rashid et al.** [29] incorporated GANs to generate synthetic lesion images, aiding in model training. The classification system outperformed standard CNNs like ResNet-50. Their GAN-augmented approach addressed dataset imbalance. It also boosted performance in minority classes like melanoma. The study showed potential in data-scarce medical scenarios. (2021)​

**Bisla et al.** [30] focused on refining ResNet-50 through data cleaning and GAN-based augmentation. Their pipeline improved classification across melanoma, nevus, and seborrheic keratosis. The GAN generated diverse and realistic lesions. The model was fine-tuned for three-class dermoscopic classification. It emphasized the importance of high-quality and diverse training data. (2021)​

**Anand et al.** [31] designed a pipeline merging U-Net for segmentation with CNN-based classification. The model first isolated the lesion area and then categorized it. This two-stage method improved overall diagnostic accuracy. It addressed both localization and classification in a unified framework. The system showed strong results on public skin cancer datasets. (2023)​

**Duan et al.** [32] designed a deep CNN framework for early melanoma recognition. They trained and validated the model on the PH2 and ISIC datasets. The model achieved high accuracy and sensitivity. Their study focused on enhancing model robustness across image variations. It contributed to building reliable AI-assisted diagnostic tools. (2023)

**Chapter 3**

**PROPOSED METHODOLOGY/TECHNIQUES**

The proposed skin cancer detection system is implemented using a combination of image processing techniques and classical machine learning algorithm.

#### 3.1 Techniques used:

Six popular machine learning algorithms are used and assessed to guarantee strong classification performance. Every one of these algorithms has unique benefits, presumptions, and behaviors that add up to a comparative evaluation of performance. A thorough rundown of the classifiers utilized is provided below:

**3.1.1 Logistic Regression (LR):**

A statistical model that models the likelihood of a binary outcome using a logistic function. It establishes a linear decision boundary to distinguish between benign and malignant lesions in the context of skin cancer detection. Particularly when the data is linearly separable, it is straightforward but efficient.

**3.1.2 k-Nearest Neighbors (k-NN):**

Data points are grouped by the majority label among their k nearest neighbors in the feature space using a non-parametric, instance-based learning algorithm known as k-Nearest Neighbors (k-NN). It performs best when the class distribution is locally consistent and is sensitive to the choice of k and distance metric.

**3.1.3 Naive Bayes (NB):**

Naive Bayes (NB), a probabilistic classifier based on Bayes' theorem, makes the assumption that, given the class label, every feature is conditionally independent. Despite its strong independence assumption, it is often very efficient and performs well in practice, particularly for high-dimensional data.

**3.1.4 Decision Tree (DT):**

A tree structure that resembles a flowchart, with internal nodes standing in for feature tests, branches for test results, and leaf nodes for class labels. It selects features automatically during training and is simple to understand and visualize.

**3.1.5 Random Forest (RF):**

A technique for ensemble learning that constructs several decision trees and uses majority voting to aggregate their predictions. This improves generalization and decreases overfitting. Even when a lot of features are involved, it performs well and is resilient to noise.

**3.1.6 Support Vector Machine (SVM):**

An effective supervised learning algorithm that creates the best hyperplane to divide data points from various classes in a high-dimensional space. SVM can efficiently handle both linear and non-linear classification tasks by utilizing kernel functions

A labeled dataset of dermoscopic images with extracted features like color, texture, shape, and asymmetry is used to train and evaluate each algorithm. To evaluate each classifier's performance, evaluation metrics such as accuracy, precision, recall, and F1-score are calculated.

**3.2 Proposed Method**

The skin cancer detection method consists of different steps, such as selecting inputs, extracting features from the inputs, and classification of skin cells as bening or malignant cell. All the steps of the skin cancer classification method are shown in the flowchart in Fig. 3.1. First, the cancer images are obtained from skin cancer patients. Then the images are processed using gray-scale conversion and normalization then features are extracted. After that the training skin cancer classification modules are designed. Subsequently, the training modules are tested with new images to find the cancer types.

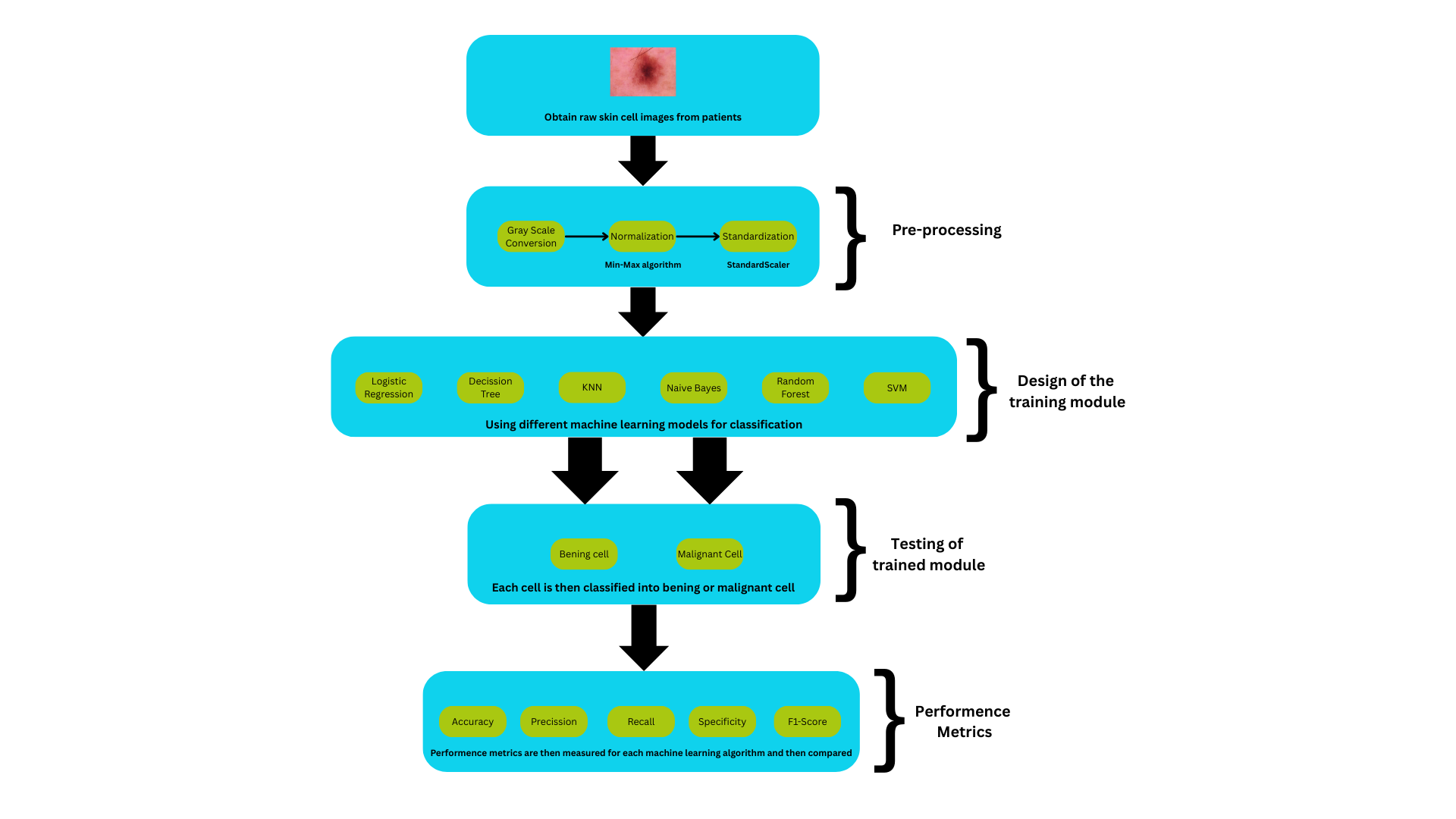


Fig1. Flowchart of the proposed method

**3.3 Materials Used**

The **HAM10000 dataset** is a benchmark resource used for developing automated skin cancer detection methods. HAM10000 offers 10,015 dermatoscopic images of various pigmented skin lesions, including melanoma and bening cells. The images are labeled with lesion types, the mask it provides and the types of cancer cells,i.e Melanoma,Melanocytic nevi, Basal cell carcinoma, Actinic keratosis, Benign keratosis-like lesion, Dermatofibroma and Vascular lesions.This diversity supports robust machine learning model training and evaluation.

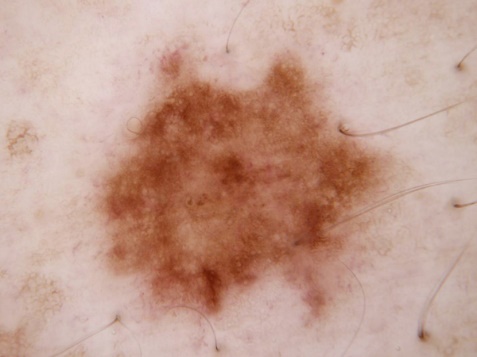
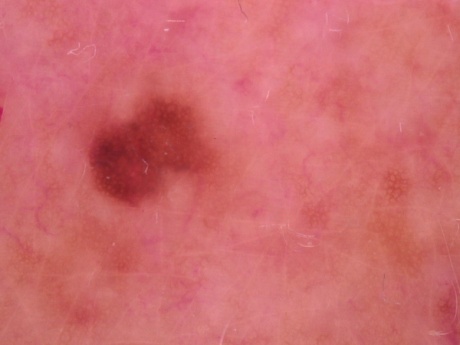
  

Fig2. Skin images extracted from HAM10000 dataset

**3.3 Pre-processing**

**3.3.1 Normalizaion**

After converting the skin images in grayscale, we normalize pixel values to the range[0, 1] by dividing by 255.0 for standard image processing purpose.This helps reduce variation across images and stabilizes model training.

**Formula**:

**3.3.2 Standardization:**

For models sensitive to scale (like SVM), we apply standardization after normalization. Each image's pixel values are transformed to have zero mean and unit variance.

**Formula**:

Where:

* μ is the mean pixel value
* σ is the standard deviation

**Chapter - 4**

## RESULT & DISCUSSION

The results of the proposed Skin Cancer Detection Systems are analyzed based on accuracy, real-time performance, and reliability across various image samples. The system integrates image pre-processing, feature extraction techniques, and a machine learning-based classifier to effectively detect and classify skin lesions.

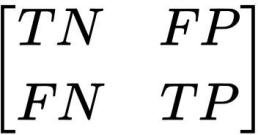
* 1. **Model Performance Evaluation :**

The performance of the skin cancer detection models was evaluated using key metrics such as accuracy, precision, specificity, recall, F1-score, confusion matrix, and ROC curve. The models tested include:

* K-Nearest Neighbors (KNN)
* Decision Tree
* Naïve Bayes
* Support Vector Machine (SVM)
* Random Forest
* Logistic Regression

#### Confusion Matrix Analysis

The confusion matrix provides insights into the model's classification accuracy by comparing actual and predicted labels.



where:

TP (True Positives): Correctly classified skin cancer cases.

TN (True Negatives): Correctly classified non-skin cancer cases.

FP (False Positives): Incorrectly classified non-skin cancer cases as drowsy.

FN (False Negatives): Incorrectly classified skin cancer cases as non-skin cancer.

#### Other Metrics

#### In classification, the performance matrices used to validate a method are recall, specificity, accuracy,

#### precision and F1-score. Recall is used to determine the method’s ability to correctly detect the cancer.

#### Specificity is the ability to reject a cancer type. Accuracy is used to determine the correctness of the

#### method. Precision tells you how many of the predicted cancerous cells are actually having cancer.An

#### d F1-score is the harmonic mean of precision and recall.This metrices have been calculated using the

#### equation as given below.

#### 

#### 

#### 

#### 

This performance metrics are then compared for each of the six machine learning algorithms as given below in table 4.1.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Algorithms used | Accuracy(in %) | Recall(in %) | Specificity(in %) | Precision(in %) | F1 Score(in %) |
| Logistic Regression | 41.34% | 80.04% | 31.95% | 22.19% | 34.74% |
| k-Nearest Neighbour  (k-NN) | 39.15% | 84.34% | 28.19% | 22.16% | 35.10% |
| Support Vector Machine(SVM) | 46.38% | 87.00% | 36.53% | 24.94% | 38.77% |
| Naïve Bayes | 66.28% | 40.89% | 72.43% | 26.45% | 32.12% |
| Decision Tree | 46.27% | 76.71% | 38.89% | 23.33% | 35.78% |
| Random Forest | 32.99% | 94.27% | 18.13% | 21.82% | 35.44% |

Table 4.1 Comparison of performance among other models

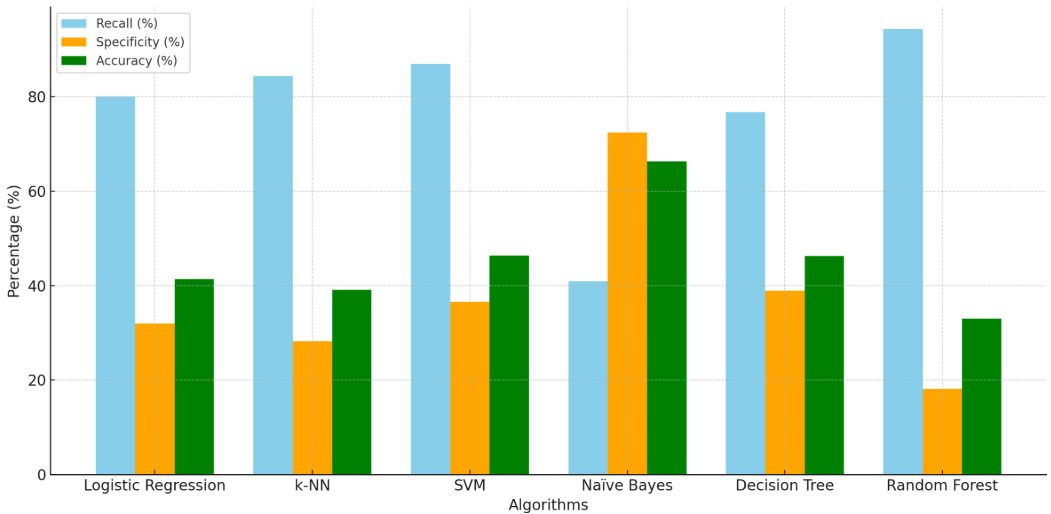
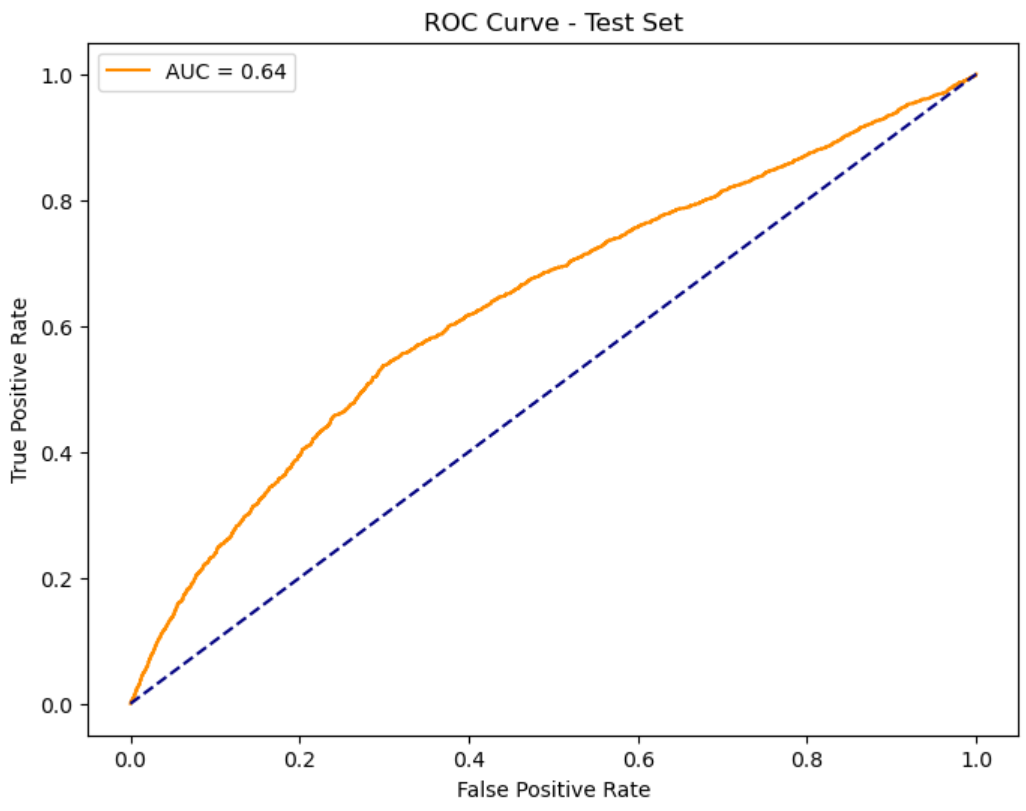
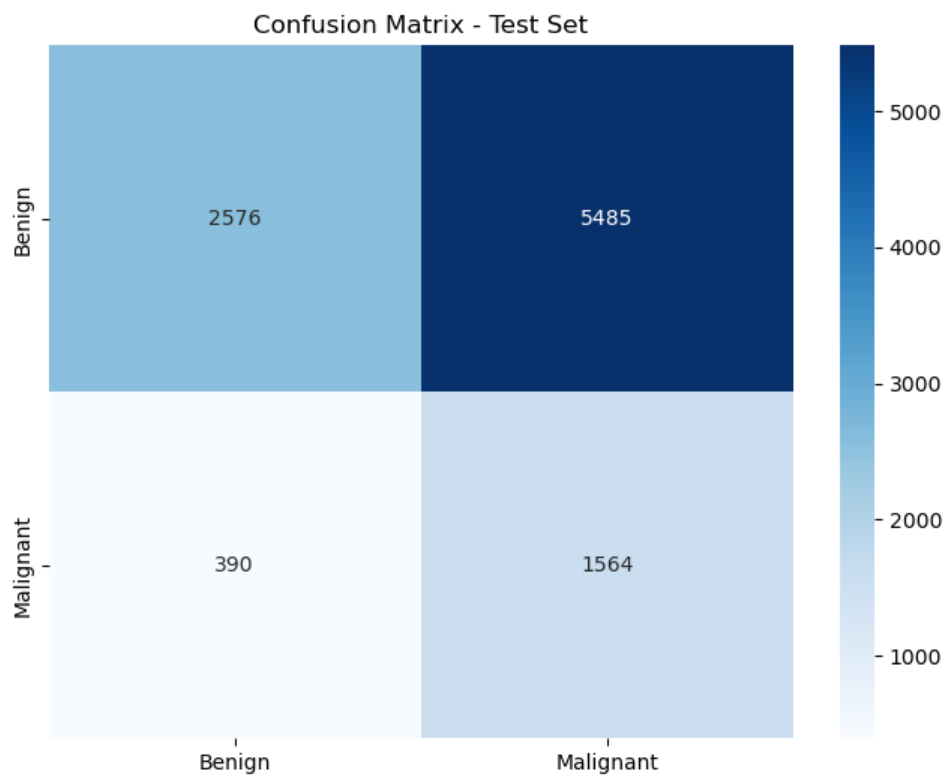


Fig3. Performance of three metrics used

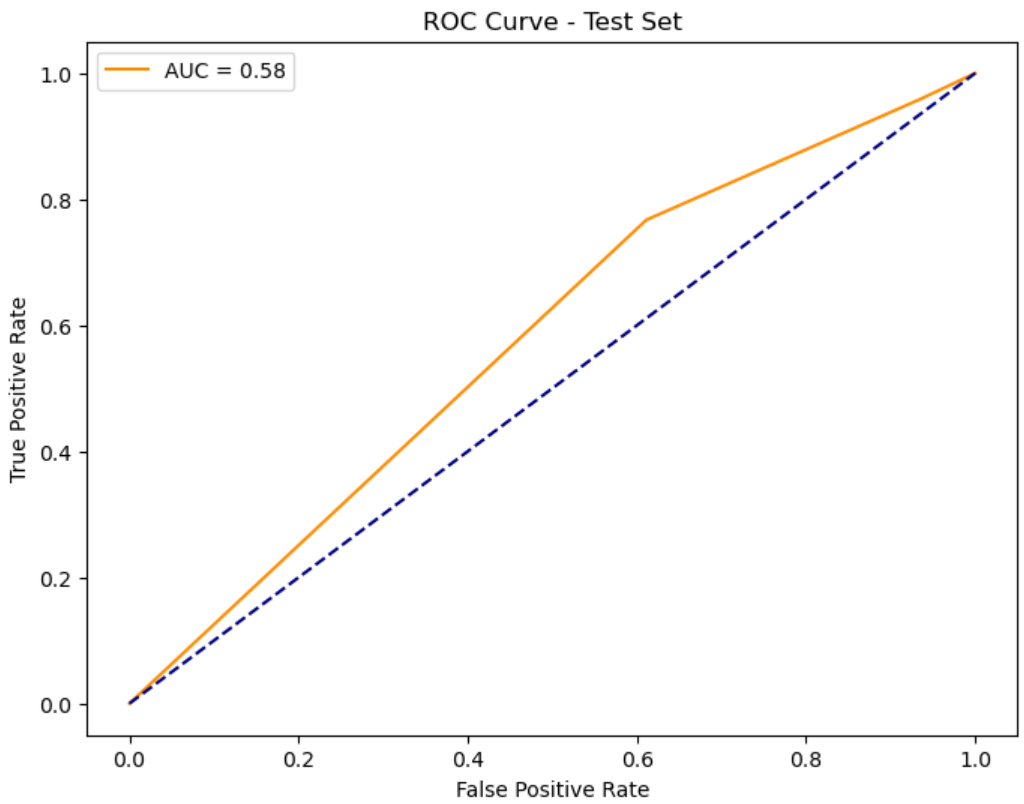
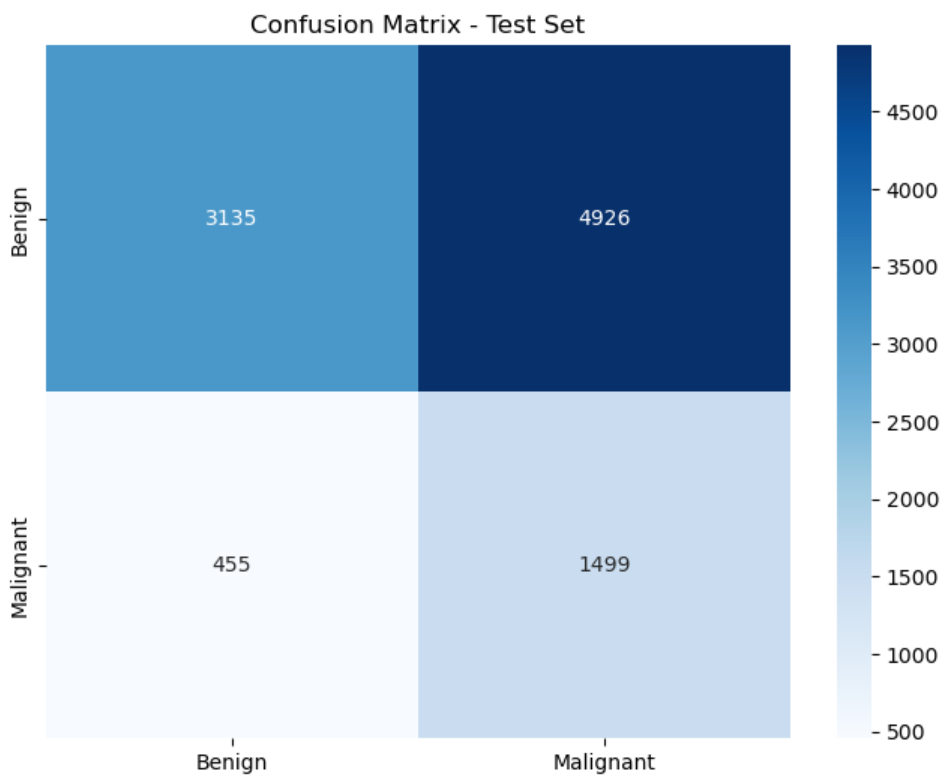
**4.1.3 Performanc Analysis of Logistic Regression:**

With an AUC of 0.64, the Logistic Regression model for skin cancer detection's overall discriminatory power is restricted, meaning it performs just marginally better than chance. It is highly successful in identifying patients who genuinely have skin cancer, as evidenced by its high recall of 80.04%. However, a relatively low Specificity of 31.95% indicates that the model commonly misclassifies healthy people as having cancer, resulting in a significant number of false alarms, severely undermining its strength. Its poor precision of 22.19%, which indicates that the model is wrong nearly four times out of five when predicting cancer, further emphasises this problem. This imbalance is confirmed by the poor overall Accuracy (41.34%) and F1 Score (34.74%), which eventually make the model's performance unsuitable for dependable real-world application because of the high percentage of false positives despite its sensitivity to real-world scenarios.

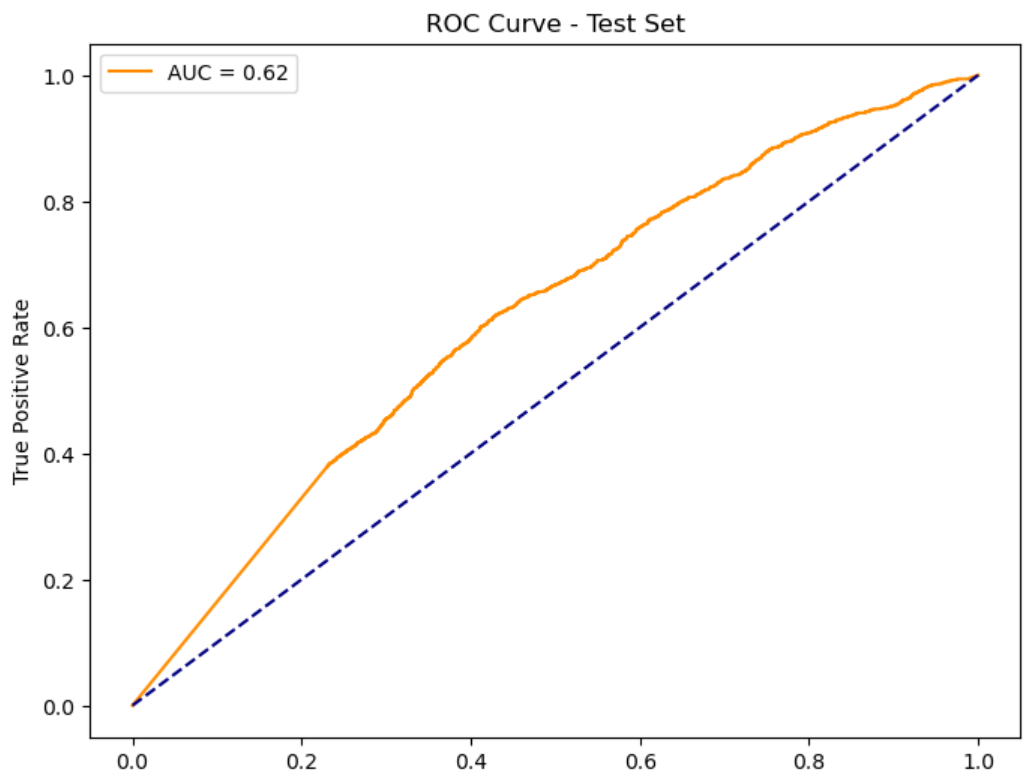
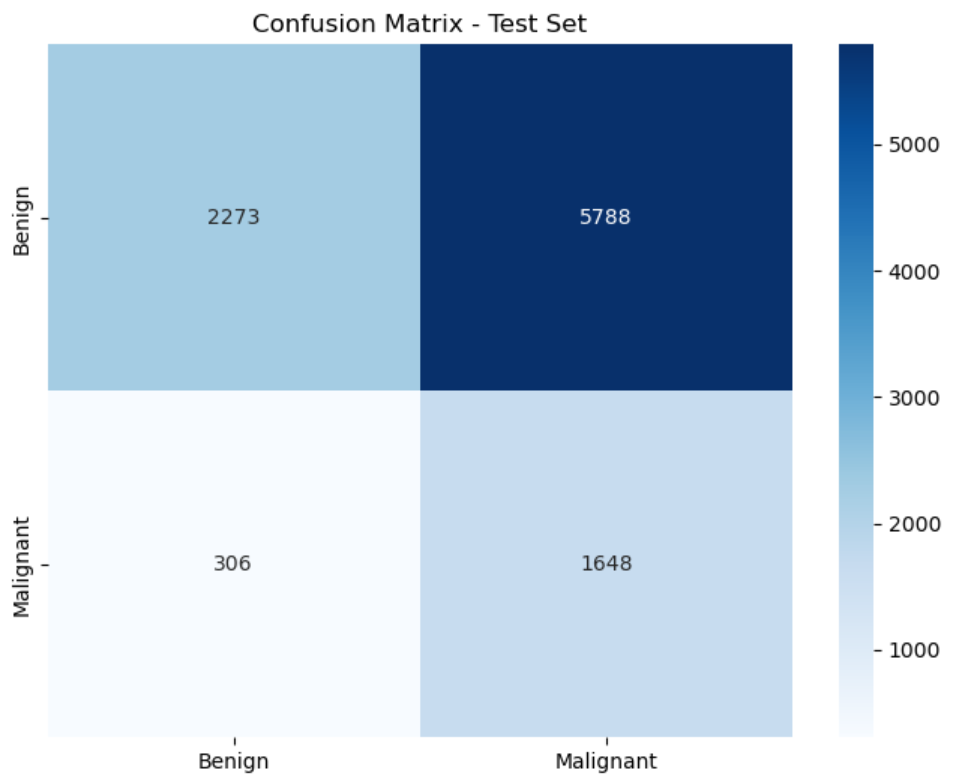
 

**4.1.4 Performance Analysis of Decision Tree:**

the Decision Tree algorithm displays low overall performance for skin cancer diagnosis, expressed by an AUC of just 0.58, which is just marginally better than random guessing. Although it can detect more than three-quarters of real cancer cases, as seen by its comparatively high recall of 76.71%, this comes with a substantial expense. The Specificity is poor at 38.89%, suggesting the model struggles to properly identify healthy patients, resulting in a substantial number of false positives. The extremely low Precision of 23.33%, which indicates that the model is wrong more than 75% of the time when it predicts cancer, further supports this. The F1 Score (35.78%) and low Accuracy (46.27%) highlight the model's inadequate sensitivity-specificity balance, making it unreliable for clinical application due to the high rate of false alarms it generates.

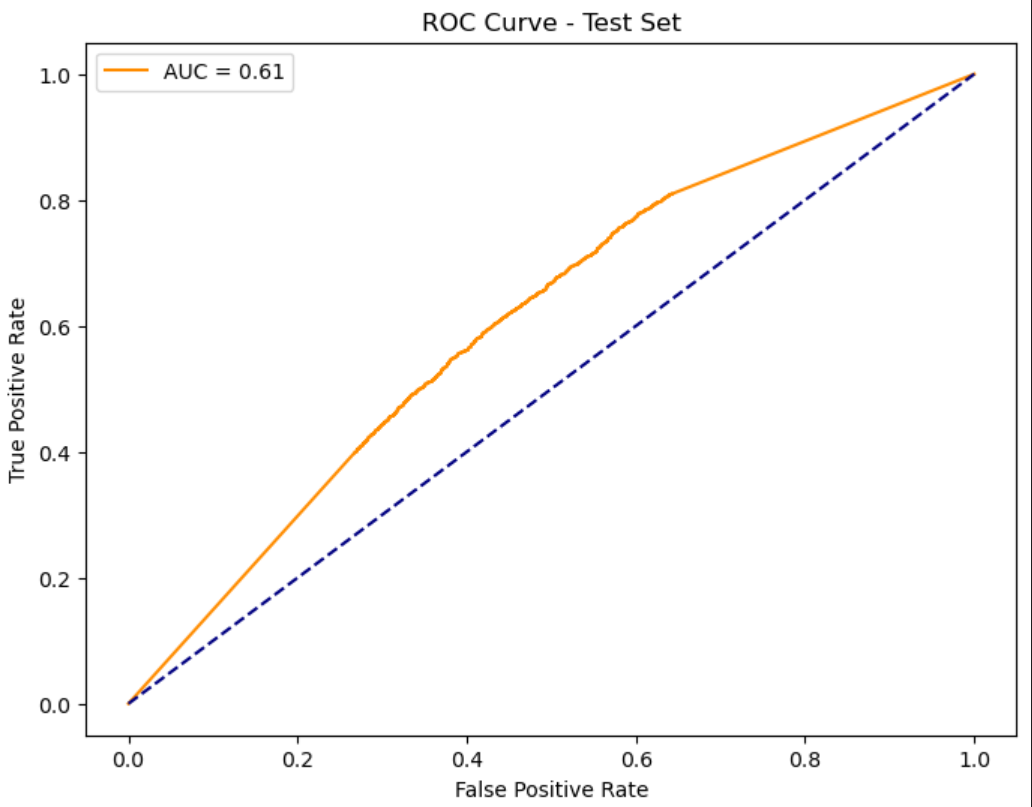
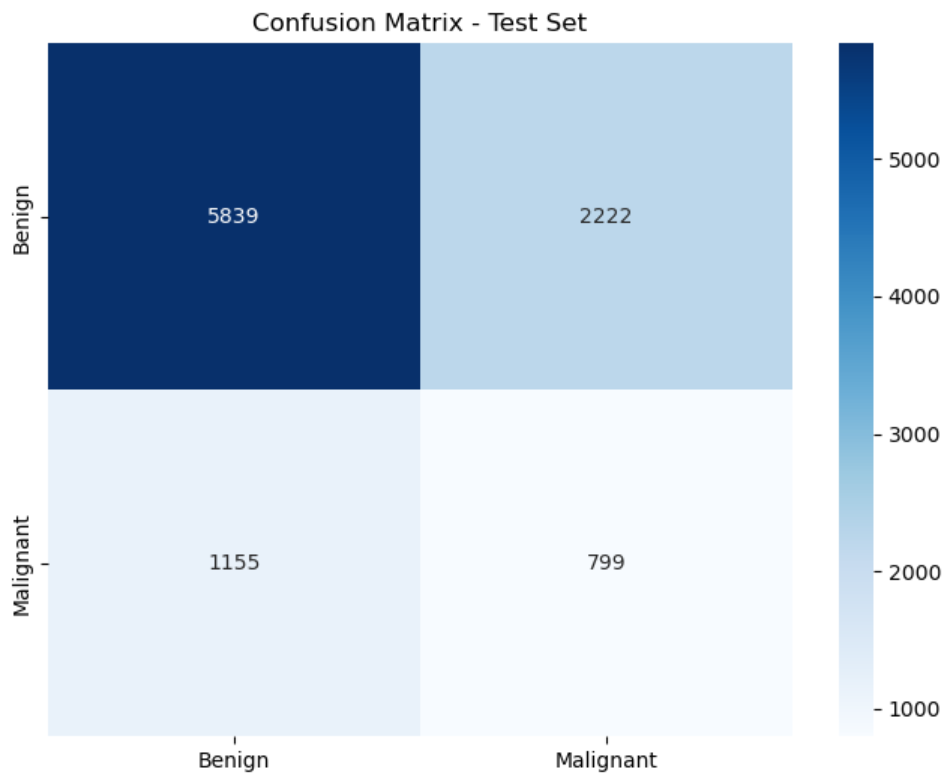
 

**4.1.5 Performance Analysis of k-Nearest Neighbor:**  
  
When it comes to identifying skin cancer, the k-Nearest Neighbour (k-NN) algorithm performs poorly overall, just marginally better than chance in differentiating between malignant and non-cancerous cases. Its extremely high recall of 84.34%, which shows that it correctly detects the vast majority of real positive skin cancer cases, is its most noteworthy feature. Nevertheless, this high sensitivity is significantly offset by a very low specificity of only 28.19%, indicating that the model produces a large number of false positives and is very bad at accurately identifying healthy individuals. This low specificity directly contributes to an extremely low precision of 22.16%, meaning that k-NN is incorrect almost 78% of the time when it predicts cancer. The resulting overall Accuracy is also very low at 39.15%, and the F1 Score of 35.10% confirms the substantial imbalance, making the k-NN model unsuitable for practical deployment due to the excessive false alarm rate despite its ability to detect true cases.

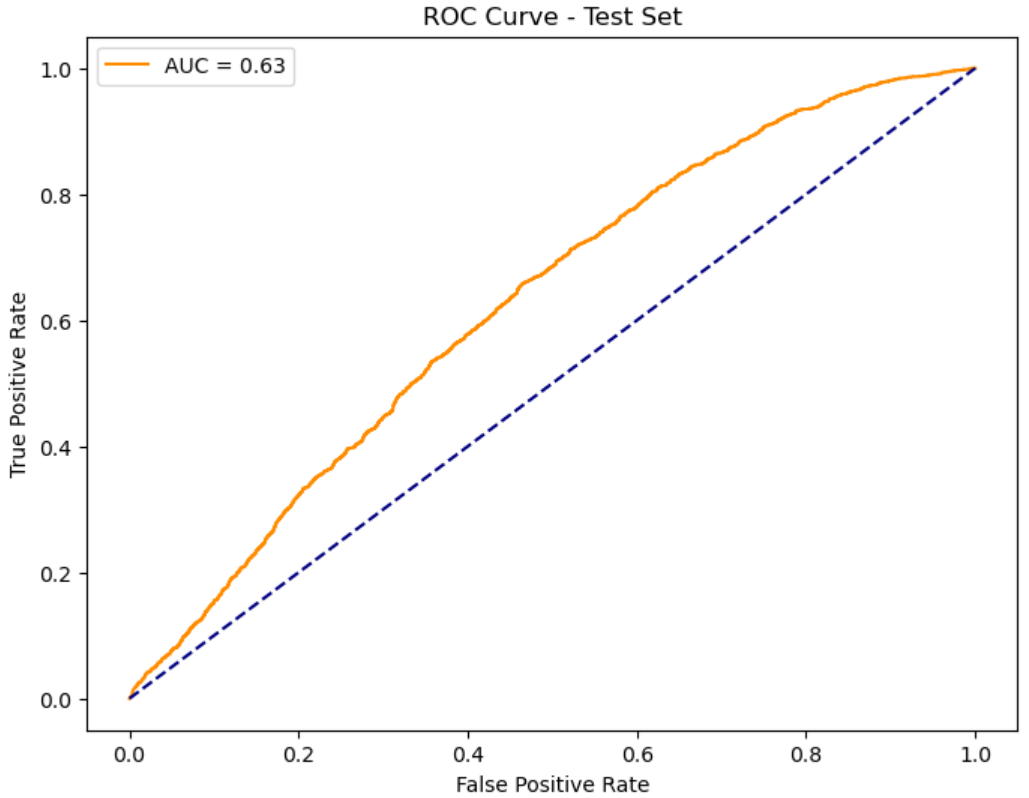
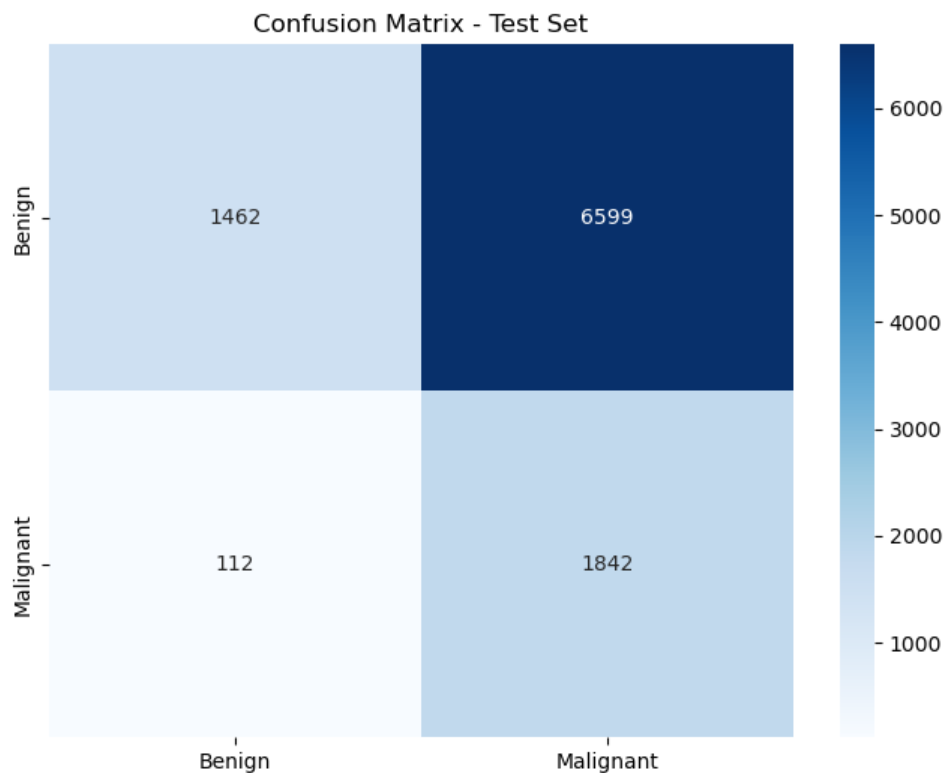
**4.1.6 Performance Analysis of Naïve Bayes:**

The Naïve Bayes algorithm performs marginally better than random chance in terms of overall classification capacity for skin cancer diagnosis. Its comparatively greater Specificity (72.43%) sets it apart from the other models. This means that it is more accurate in identifying individuals who do not have cancer, resulting in fewer false positives than genuine negatives. This comes at a terrible cost, though, since the model's recall rate of 40.89% is abnormally low, meaning it misses over 60% of real skin cancer cases, which is extremely risky in a medical setting. Because of the fundamental failure in sensitivity, this measure is deceptive even though it has the greatest accuracy (66.28%) of the algorithms given. The Precision remains low at 26.45%, and the very low F1 Score (32.12%) reflects the extreme imbalance and poor performance caused by missing so many positive cases, rendering Naïve Bayes unsuitable for this critical detection task.

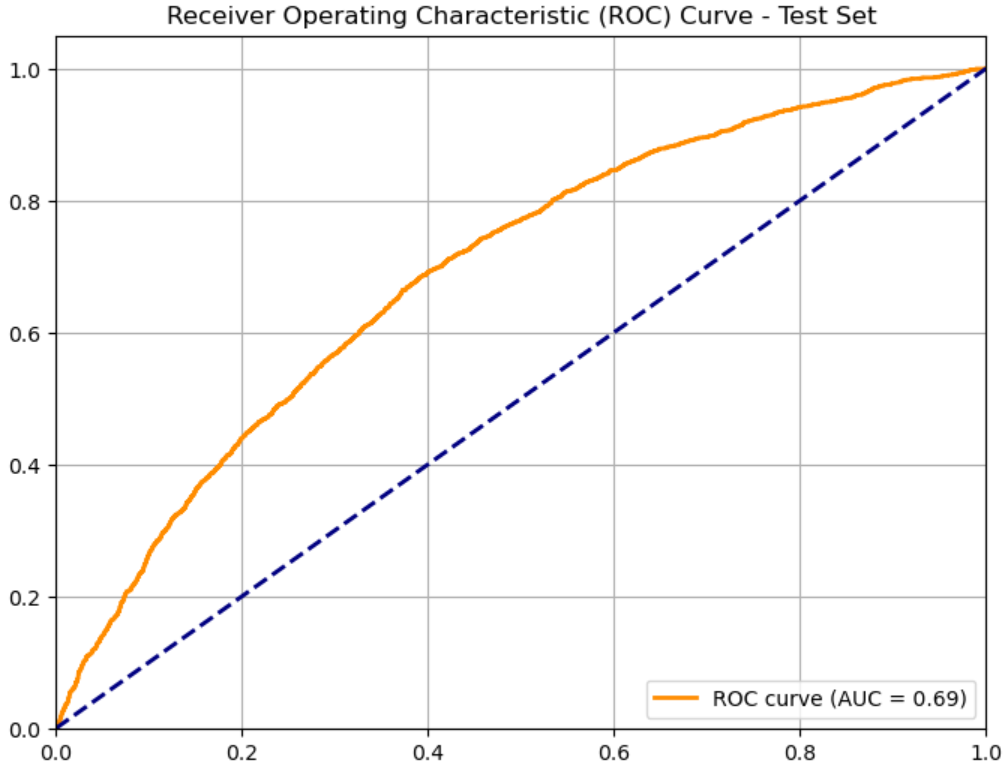
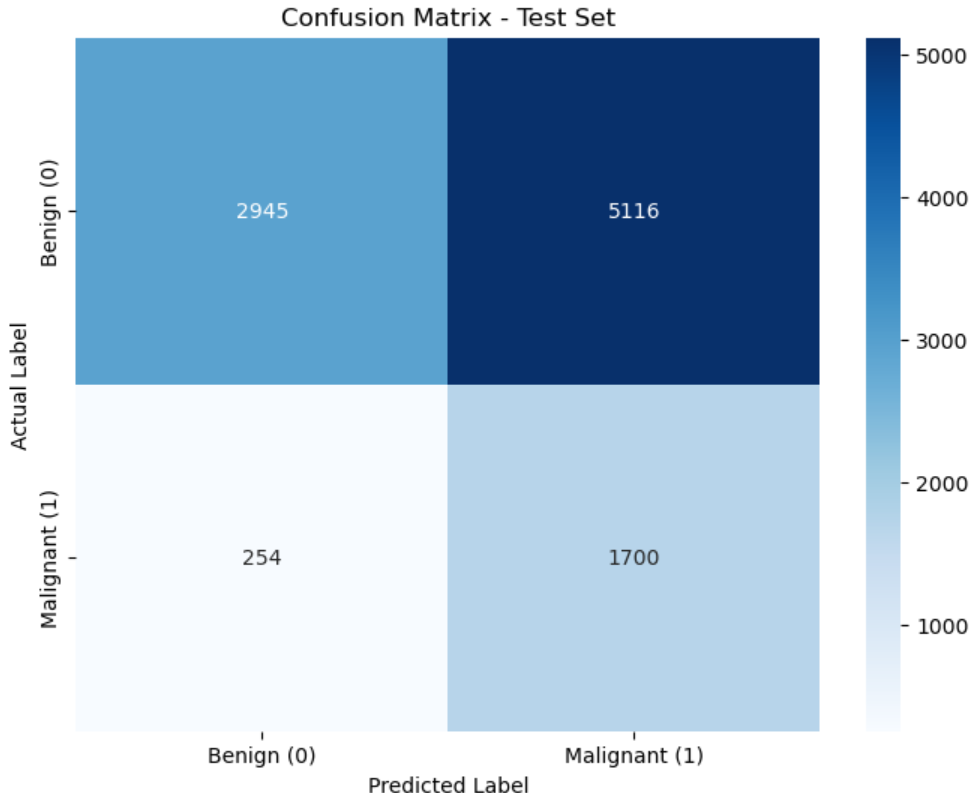
**4.1.7 Performance Analysis of Random Forest:**

With an AUC of 0.63 as indicated by the ROC curve, the Random Forest algorithm performs just marginally better than random chance in distinguishing between skin diseases that are malignant and those that are not. Its profile is characterised by an extraordinary trade-off: it is extraordinarily successful at recognising almost all real cancer patients, as seen by its exceptionally high Recall of 94.27%. This is seriously compromised, though, by the model's incredibly low Specificity of only 18.13 percent, which causes an overwhelming amount of false positives by mistakenly classifying the great majority (more than 80%) of healthy individuals as having cancer. As a result, the precision is extremely poor at 21.82%, meaning that the model is wrong around four times out of five when it predicts cancer. This results in a very low overall Accuracy of 32.99% and a low F1 Score of 35.44%, highlighting the severe imbalance and making the model highly impractical for clinical use due to the massive burden of false alarms, despite its high sensitivity.

**4.1.8 Performance Analysis of Support Vector Machine (SVM):**

The Support Vector Machine (SVM) approach outperforms a number of other models examined, but it still falls short of high reliability in its ability to discriminate between skin lesions that are malignant and those that are not. Its most notable strength is its extremely high recall of 87.00%, which shows that it correctly detects the vast majority of real instances of skin cancer. However, a low specificity of 36.53% severely undermines this high sensitivity, causing the model to often misclassify healthy individuals, resulting in a considerable number of false positives. As a result, the precision is rather poor at 24.94%, meaning that around 75% of the time, SVM is wrong when predicting cancer. Its moderate accuracy of 46.38% and its F1 Score of 38.77%, which is the highest among the models assessed, indicate a little better balance between precision and recall than the others, but it is still poor overall. As a result, the SVM's practical value for accurate skin cancer detection is limited by a large false positive rate, even if it has the best F1 score and good recall.

**4.2 Discussion:**

When six different machine learning algorithms—Logistic Regression, k-NN, SVM, Naïve Bayes, Decision Tree, and Random Forest—are evaluated for the identification of skin cancer, the results show consistently subpar performance. A distinct and troubling pattern shows up: the majority of algorithms get high recall (sensitivity), successfully detecting a significant portion of real cancer cases (ranging from 77% to 94% for all but Naïve Bayes). But this comes at the terrible cost of having very low specificity, which leads to very poor precision (all below 27%) and, as a result, a large false positive rate. This severe trade-off is best illustrated by Random Forest, which has the lowest Specificity (18.13%) and the highest Recall (94.27%).

Naïve Bayes, on the other hand, would miss about 60% of tumours because to its unacceptable low recall (40.89%) and highest specificity (72.43%). With the greatest F1 score (38.77%) and a comparatively high AUC (0.69), SVM exhibits the marginally best balance; yet, like the others, its performance is typified by poor overall accuracy and discriminating power (AUCs range 0.58-0.69). In conclusion, the essential imbalance that results in either dangerously high levels of missed detections or excessive false alarms is the main reason why none of the assessed models exhibit the dependability required for clinical usage in their current condition.

**Chapter 5**

**5.1 Conclusion**  
Six machine learning algorithms for skin cancer diagnosis were examined in this work using the provided dataset and metrics: Random Forest, Naïve Bayes, Support Vector Machine, k-Nearest Neighbour, Logistic Regression, and Decision Tree. None of them met the dependability requirements for clinical application, according to the data.   
Many false positives resulted from the high recall (76.71% to 94.27%), very low specificity (18.13% to 38.89%), and poor accuracy (all below 27%) of the majority of models (K-NN, SVM, Decision Tree, Random Forest, and Logistic Regression). Clinically, this would make patients anxious and lead to needless biopsies. On the other hand, despite its seeming high accuracy (66.28%), Naïve Bayes obtained a greater specificity (72.43%) but a relatively poor recall (40.89%), putting many undetected cancer cases at danger.

Weak discrimination and an imbalance between recall and accuracy were further supported by low F1 scores (all below 39%) and AUC values (0.58–0.69).   
The underperformance is probably caused by job complexity, class imbalance, insufficient feature representation, or dataset restrictions. To attain clinically feasible performance, future advancements should investigate deep learning models such as CNNs, improved class imbalance management, sophisticated feature engineering, hyperparameter tweaking, and bigger, more varied datasets.

**5.2 FUTURE WORK**

While the current implementation focuses on image-based skin cancer detection using classical machine learning models and image processing techniques, several future enhancements can be explored to improve system accuracy, scalability, and usability:

**Deep Learning Integration:** Incorporating advanced deep learning models such as Convolutional Neural Networks (CNNs), MobileNetV2, or EfficientNet could significantly enhance classification accuracy by learning complex lesion patterns directly from raw images.

**IoT-Enabled Remote Diagnosis:** Integration with IoT-based medical devices and mobile health platforms can enable real-time remote screening and consultation, especially beneficial for rural or underserved areas lacking dermatological facilities.

**Cloud-Based Data Management:** Implementing a cloud-based architecture for storing patient records, image data, and classification results can facilitate large-scale deployments, longitudinal tracking, and centralized medical data access for healthcare professionals.

**Mobile and Embedded Deployment:** Porting the system to mobile platforms or lightweight edge devices such as Raspberry Pi, Jetson Nano, or Android-based apps would make the solution portable and accessible for real-world usage, including on-the-go diagnostics.

**Personalized Risk Assessment:** Developing patient-specific models that consider factors like age, skin type, and lesion history could improve diagnostic precision and reduce misclassification rates.

**User-Friendly Interface and Reporting:** Enhancing the user interface with interactive visualizations, easy navigation, and downloadable reports can improve usability for both medical practitioners and patients.

By implementing these enhancements, the proposed system can evolve into a comprehensive, intelligent diagnostic tool, supporting early skin cancer detection and contributing meaningfully to accessible, technology-driven healthcare.

**Chapter - 6**

## REFERENCES

1. Healthline, "Basal Cell Carcinoma vs. Squamous Cell Carcinoma," *Healthline*, [Online]. Available: https://www.healthline.com/health/skin-cancer/basal-cell-carcinoma-vs-squamous-cell-carcinoma.

2. Verywell Health, "The ABCDEs of Melanoma," *Verywell Health*, [Online]. Available: https://www.verywellhealth.com/the-abcds-of-melanoma-1069472.

3. GQ, "Why Does Melanoma Disproportionately Affect Men?" *GQ*, [Online]. Available: https://www.gq.com/story/why-does-melanoma-disproportionately-affect-men.

4. M. J. Gómez, R. Hornero, and D. Álvarez, "Automatic computer diagnosis of melanoma using multiresolution texture descriptors," *BioMedical Engineering OnLine*, vol. 12, no. 1, 2013. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3885227/>.

**5.** World Cancer Research Fund International, "Skin cancer statistics," [Online]. Available: <https://www.wcrf.org/preventing-cancer/cancer-statistics/skin-cancer-statistics/>.

**6.** Sun Skin Clinic, "The Importance of Early Detection in Skin Cancer: Increasing Survival Rates," [Online]. Available: <https://sunskinclinic.com.au/the-importance-of-early-detection-in-skin-cancer-increasing-survival-rates/>.

**7.** Cochrane, "How accurate is dermoscopy compared to visual inspection of the skin in diagnosing skin cancer (melanoma) in adults?" [Online]. Available: <https://www.cochrane.org/CD011902/SKIN_how-accurate-dermoscopy-compared-visual-inspection-skin-diagnosing-skin-cancer-melanoma-adults>.

**8.** Stanford Medicine, "AI improves accuracy of skin cancer diagnoses in Stanford Medicine-led study," [Online]. Available: <https://med.stanford.edu/news/all-news/2024/04/ai-skin-diagnosis.html>.

**9.** The American Journal of Managed Care, "AI-Based Smartphone App Proves Reliable in Diagnosis of Melanoma," [Online]. Available: <https://www.ajmc.com/view/ai-based-smartphone-app-proves-reliable-in-diagnosis-of-melanoma>. [Accessed: Apr. 24, 2025].​[AJMC](https://www.ajmc.com/view/ai-based-smartphone-app-proves-reliable-in-diagnosis-of-melanoma?utm_source=chatgpt.com)

**10.** Annals of Oncology, "Artificial intelligence for melanoma diagnosis: how can we deliver on the promise?" [Online]. Available: <https://www.annalsofoncology.org/article/S0923-7534(19)34114-6/fulltext>.

11. M. Kassem, K. Hosny, and M. M. Damaševičius, "Machine learning and deep learning methods for skin lesion classification and diagnosis: A systematic review," *Diagnostics*, vol. 10, no. 8, p. 580, 2020.

12. M. Bistroń and A. Piotrowski, "A comparative analysis of machine learning algorithms for skin cancer diagnosis," *Applied Sciences*, vol. 12, no. 22, p. 11390, 2022.

13. H. Nguyen, T. Nguyen, and H. Nguyen, "Comparative analysis of machine learning models for melanoma diagnosis with dermoscopic images," *Journal of Imaging*, vol. 10, no. 1, p. 34, 2024.

14. E. Nasr-Esfahani, S. Samavi, N. Karimi, S. Soroushmehr, K. Ward, and K. Najarian, "Melanoma detection by analysis of clinical images using convolutional neural network," *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 1373–1376, 2016.

15. A. Esteva, B. Kuprel, R. A. Novoa, et al., "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, 2017.

16. L. Yu, H. Chen, Q. Dou, J. Qin, and P.-A. Heng, "Automated melanoma recognition in dermoscopy images via very deep residual networks," *IEEE Transactions on Medical Imaging*, vol. 36, no. 4, pp. 994–1004, Apr. 2017.

17. J. Zhang, J. Xie, Y. Xia, and C.-S. Lee, "Attention residual learning for skin lesion classification," *IEEE Transactions on Medical Imaging*, vol. 38, no. 9, pp. 2092–2103, Sept. 2019.

18. T. Brinker, A. Hekler, F. Enk, et al., "Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task," *European Journal of Cancer*, vol. 113, pp. 47–54, 2019.

19. P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Scientific Data*, vol. 5, no. 1, p. 180161, 2018.

20. M. Mahbod, R. Schaefer, and G. E. Santillan, "Effects of image resolution on deep learning performance for skin lesion classification," *IEEE Access*, vol. 8, pp. 143615–143627, 2020.

21. P. Gessert, M. Nielsen, M. Shaikh, F. Werner, A. Schlaefer, and T. B. Huber, "Skin lesion classification using ensembles of multi-resolution EfficientNets and meta learning," *Medical Image Analysis*, vol. 67, p. 101789, 2021.

22. J. Amin, M. Sharif, M. Yasmin, and S. L. Fernandes, "A distinctive approach in skin lesion detection and classification using fine-tuned deep neural network," *Pattern Recognition Letters*, vol. 138, pp. 370–377, 2020.

23. S. Bakkouri and K. Afdel, "Skin lesion diagnosis system based on a hybrid deep learning approach," *Multimedia Tools and Applications*, vol. 80, pp. 12773–12796, 2021.

24. X. Wei, Z. Yuan, and Y. Qin, "Lightweight convolutional neural networks for mobile skin cancer diagnosis," *IEEE Access*, vol. 8, pp. 109760–109770, 2020.

25. E. Kaymak, A. Uyar, and S. Kaymak, "Segmentation of skin lesions using fully convolutional networks," *Computerized Medical Imaging and Graphics*, vol. 88, p. 101820, 2021.

26. S. Khan, S. Islam, M. Jan, and J. Din, "Automated skin lesion detection and classification using saliency-based preprocessing and deep learning," *Sensors*, vol. 20, no. 22, p. 6604, 2020.

27. R. Jain, S. Nayyar, A. Jain, and A. Singhal, "Comparative analysis of transfer learning models for classification of skin lesion," *Materials Today: Proceedings*, vol. 47, pp. 5620–5624, 2021.

28. A. Kausar, M. R. Syed, M. Jamil, and S. A. Khan, "Ensemble deep learning approach for melanoma detection," *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, pp. 773–783, 2022.

29. M. Rashid, F. Akbar, and M. Hassan, "GAN-based synthetic image augmentation for improving skin lesion classification," *Computers in Biology and Medicine*, vol. 134, p. 104426, 2021.

30. S. Bisla, A. Singh, and R. Kumar, "Improving skin lesion classification by data cleaning and GAN-based data augmentation," *Biomedical Signal Processing and Control*, vol. 69, p. 102910, 2021.

31. S. Anand, P. Sahu, and M. Sharma, "Melanoma skin cancer detection using U-Net and CNN classification," *Biomedical Signal Processing and Control*, vol. 79, p. 104028, 2023.

32. Y. Duan, H. Xie, J. Huang, and X. Chen, "Early melanoma recognition using a deep CNN framework," *Medical & Biological Engineering & Computing*, vol. 61, pp. 127–137, 2023.