

RESEARCH ARTICLE

A model for skin cancer using combination of ensemble learning and deep learning

Mehdi Hosseinzadeh^{1,2}, Dildar Hussain³, Firas Muhammad Zeki Mahmood⁴, Farhan A. Alenizi⁵, Amirhossein Noroozi Varzeghani⁶, Parvaneh Asghari⁶, Aso Darwesh⁷, Mazhar Hussain Malik^{8*}, Sang-Woong Lee^{9*}

1 Institute of Research and Development, Duy Tan University, Da Nang, Vietnam, **2** School of Medicine and Pharmacy, Duy Tan University, Da Nang, Vietnam, **3** Department of AI and Data Science, Sejong University, Seoul, Republic of Korea, **4** Department of Communication and Computer Engineering, Cihan University-Erbil, Kurdistan Region, Iraq, **5** Electrical Engineering Department, College of engineering, Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia, **6** Department of Computer Engineering, Central Tehran Branch, Islamic Azad University, Tehran, Iran, **7** Department of Information Technology, University of Human Development, Sulaymaniyah, Kurdistan region of Iraq, **8** School of Computer Science and Creative Technologies College of Arts, Technology and Environment (CATE) University of the West of England Frenchay Campus, Coldharbour Lane Bristol, Bristol, United Kingdom, **9** Pattern Recognition and Machine Learning Lab, Gachon University, Seongnamdaero, Sujeonggu, Seongnam, Republic of Korea

☞ These authors contributed equally to this work.

* mazhar.malik@uwe.ac.uk (HM); slee@gachon.ac.kr (SWL)



OPEN ACCESS

Citation: Hosseinzadeh M, Hussain D, Zeki Mahmood FM, A. Alenizi F, Varzeghani AN, Asghari P, et al. (2024) A model for skin cancer using combination of ensemble learning and deep learning. PLoS ONE 19(5): e0301275. <https://doi.org/10.1371/journal.pone.0301275>

Editor: Sathishkumar Veerappampalayam Easwaramoorthy, Sunway University, MALAYSIA

Received: January 22, 2024

Accepted: March 13, 2024

Published: May 31, 2024

Copyright: © 2024 Hosseinzadeh et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript.

Funding: This study is supported by funding from Prince Sattam bin Abdulaziz University to FAA [project number PSAU/2024/R/1445].

Competing interests: The authors have declared that no competing interests exist.

Abstract

Skin cancer has a significant impact on the lives of many individuals annually and is recognized as the most prevalent type of cancer. In the United States, an estimated annual incidence of approximately 3.5 million people receiving a diagnosis of skin cancer underscores its widespread prevalence. Furthermore, the prognosis for individuals afflicted with advancing stages of skin cancer experiences a substantial decline in survival rates. This paper is dedicated to aiding healthcare experts in distinguishing between benign and malignant skin cancer cases by employing a range of machine learning and deep learning techniques and different feature extractors and feature selectors to enhance the evaluation metrics. In this paper, different transfer learning models are employed as feature extractors, and to enhance the evaluation metrics, a feature selection layer is designed, which includes diverse techniques such as Univariate, Mutual Information, ANOVA, PCA, XGB, Lasso, Random Forest, and Variance. Among transfer models, DenseNet-201 was selected as the primary feature extractor to identify features from data. Subsequently, the Lasso method was applied for feature selection, utilizing diverse machine learning approaches such as MLP, XGB, RF, and NB. To optimize accuracy and precision, ensemble methods were employed to identify and enhance the best-performing models. The study provides accuracy and sensitivity rates of 87.72% and 92.15%, respectively.

Introduction

The unrestricted growth of skin cells, commonly triggered by exposure to UV radiation, genetic factors, and environmental influences, is a common cause of skin cancer [1]. The sickness involves numerous histological subtypes, with basal cell carcinoma (BCC), squamous cell

carcinoma (SCC), and melanoma being the three key types [2]. Basal cell carcinoma, the most prevalent type, generally presents itself as slow-growing lesions on sun-exposed areas, which are characterized by a pearly appearance and rolled edges. Scaly or crusted lesions are a common sign of squamous cell carcinoma, which can metastasize if not identified and treated early, in contrast to other types of cancer.

Despite its high potential for metastasis, melanoma is not a primary concern compared to other cases. It arises from melanocytes, the cells that produce pigment in the skin, and it is characterized by its tendency to display various morphological characteristics and irregular borders. Skin cancer stands as the prevailing and most frequently encountered form of malignancy globally [3]. Annually, there are more than 3.5 million incidences of Melanoma, Basal Cell Carcinoma, and Squamous Cell Carcinoma diagnosed, surpassing the collective occurrences of breast cancer, lung cancer, and colon cancers. Remarkably, Melanoma alone claims a new victim every 57 seconds [4].

Skin cancer detection benefits from using machine learning and deep learning due to their capacity for automated, precise, and effective analysis of skin lesions and images. They facilitate early detection, ensure consistent performance, enable data analysis, and assist in tailoring medical decisions to individuals. Furthermore, these technologies play a role in screening, prioritization, and ongoing improvement, making them well-suited for widespread implementation. Their adoption contributes to the reduction of unnecessary diagnoses and the timely identification of potential skin cancer cases. Due to the importance of proper and accurate diagnosis of the type of skin cancer, this study employed various techniques in machine learning and deep learning to enhance the precision and accuracy of distinguishing between benign and malignant cases of skin cancer.

The contributions in this study are as follows:

- Employing a spectrum of transfer learning models, such as DenseNet-201, DenseNet-121, ResNet-50, ResNet-101, ResNet-152, VGG19, and EfficientNet-B3
- Due to the high number of features after extraction, a range of feature selectors are deployed, namely, ANOVA, XGB, Lasso, PCA, Random Forest, Mutual Information, Univariate, and Variance
- Improving evaluation metrics by deploying different Ensemble techniques and enhancing the quality of the results

The rest of this paper is outlined as follows: Section 2 delves into the discussion of related works; Section 3 explains the suggested approach; Section 4 presents experimental results and discussion; in Section 5, a comparison between others' work and this paper is depicted; and lastly, section 6 encapsulates the conclusions drawn from the study.

Related work

Recently, there has been a myriad of inquiries conducted on the matter of skin cancer. At present, multiple research initiatives are progressing toward investigating how machine learning and deep learning methodologies can be added to this field. The healthcare field has seen a substantial surge in interest in computer-assisted diagnosis (CAD), which is a vital area of investigation. This sector encompasses a plethora of research that is currently being scrutinized in the present paper.

Kumar K et al. [5] developed an advanced skin cancer classification and prediction technique using augmented intelligence with the ResNet50 model on Kaggle datasets. Using the Augmented Deep Neural Networking (AuDNN) method, they extracted critical features,

identified cancer regions, and improved dataset accuracy through clustering and attribute dependency mapping. This approach achieved 93.26% accuracy in skin cancer classification. Balaha and Hassan [6] developed an automated approach using the Sparrow Search Algorithm (SpaSA) to detect, classify, and segment skin cancer. They used U-Net models and a meta-heuristic optimizer along with pre-trained CNN models on datasets. Tabrizchi et al. [7] devised an automated skin cancer detection model using dermoscopic images and an enhanced VGG-16 CNN architecture. Adla et al. [8] crafted a robust medical decision support system by implementing an optimized full-resolution convolutional network for categorizing skin lesions from dermoscopy images. They introduced a method that fine-tuned hyperparameters using a dynamic graph cut algorithm, effectively addressing segmentation challenges and refining the precision of skin cancer diagnosis. Their model achieved 97.986% accuracy in precisely categorizing skin lesions. Mridha et al. [9] developed deep learning models for skin cancer classification, addressing class imbalance issues. Using the HAM10000 dataset, they trained a CNN, achieving an 82% classification accuracy and 0.47% loss accuracy for seven types of skin cancer. Additionally, an explainable AI system using Grad-CAM and Grad-CAM++. Wu et al. [10] introduced a skin cancer classification method using discrete wavelet down-sampling for feature reconstruction, addressing issues of information loss during down-sampling. Their approach integrated a multichannel attention mechanism improving pathological feature utilization. The proposed model achieved 95.84% accuracy. Qasim Gilani et al. [11] used a neural network approach using the surrogate gradient descent method to classify over 6,000 skin lesion images. Their proposed VGG-13 model achieved 89.57% accuracy and a 90.07% F1 score. Huang et al. [12] investigated using hyperspectral imaging (HSI) to detect skin cancer lesions, utilizing the ISIC dataset to train and test models for basal cell carcinoma (BCC), squamous cell carcinoma (SCC), and seborrheic keratosis (SK). They applied the YOLO version 5 to train the model and compared the performance of HSI and RGB classification models. The study revealed that the HSI model showed a 7.5% increase in overall recall rate (0.722 to 0.794) compared to the RGB model, particularly in capturing SCC features, indicating the potential for HSI to improve skin cancer detection. Teodoro et al. [13] addressed the challenges in early diagnosis of skin cancer due to the similarity of symptoms with other diseases, leading to errors in diagnosis. They proposed an approach using EfficientAttentionNet, a CNN architecture, for early detection of melanoma and non-melanoma skin lesions. Their methodology involved several stages: pre-processing skin images, using GAN to balance sample numbers, creating masks for regions of interest via a U-net model, and training EfficientAttentionNet with a mask-based attention mechanism. S M et al. [14] developed a Deep Convolutional Neural Network (DCNN) to classify skin cancer types from dermoscopic images. Using ISIC-2019 and ISIC-2020 datasets, they addressed image resolution differences and class imbalances through augmentation and metadata use. Leveraging the EfficientNet architecture with transfer learning and the ranger optimizer, they achieved superior skin lesion classification results, obtaining an AUC-ROC score of 0.9681 after fine-tuning EfficientNet-B6. The paper [15] introduced deep transfer learning for early skin cancer detection, leveraging pre-trained convolutional neural network models. It employed data augmentation techniques to enhance model robustness and prevent overfitting on MODE-NODE and ISIC skin lesion datasets. Through empirical analysis, vgg19 was identified as the most suitable model, achieving a testing accuracy of 98.8%. Comparative results with existing methods showed vgg19's superiority in accuracy and efficiency, trained on significantly fewer images. These findings highlighted the efficacy of data augmentation in improving detection while reducing resource consumption. Khan et al. [16] presented a fully automated computer-aided diagnosis (CAD) system for detecting malignant melanoma, a deadly form of skin cancer. It utilized a deep learning framework, including pre-processing, lesion segmentation with MASK-RCNN, and feature

extraction with DenseNet. The system achieved accuracy on validation datasets ISBI2016, ISBI2017, and HAM10000. MASK-RCNN also achieved high accuracy on ISBI2016 and ISBI2017. The paper provided a comparison with other state-of-the-art methods, demonstrating the effectiveness of the proposed framework. [17] developed a method for skin cancer detection and recognition using a combination of deep learning and an iteration-controlled Newton-Raphson (IcNR) based feature selection method. It followed three primary steps: lesion localization using Faster R-CNN with a contrast stretching approach based on the bee colony method (ABC), deep feature extraction using DenseNet201, and feature selection with IcNR. The selected features were then used for classification with multilayered feed-forward neural networks. Testing on ISBI2016 and ISBI2017 datasets achieved accuracies of 94.5% and 93.4%, respectively. Results demonstrated the proposed technique's superior accuracy and efficiency compared to existing methods. Khan et al. [18] proposed a novel deep-learning framework for lesion segmentation and classification. Mask R-CNN was implemented for segmentation, utilizing Resnet50 and feature pyramid network (FPN) as a backbone. Fully connected layers generated final masks. For classification, a 24-layered convolutional neural network was employed, activating based on visualized higher features. The method was validated on PH2, ISBI2016, and ISIC2017 datasets for segmentation and HAM10000 for classification, demonstrating superior performance with high sensitivity (85.57%), precision (87.01%), F1-Score (86.28%), and accuracy (86.5%) compared to existing techniques. To have better overview of all, we put them in Table 1.

Proposed methods

This section provides a comprehensive elucidation of the research approach employed for the diagnosis of skin cancer, encompassing all associated elements. It contains the preprocessing phase, delineates the techniques utilized for feature extraction, and discusses the algorithms applied. A condensed overview of these employed methods is visually represented in Fig 1.

A summary of the applied method is illustrated in Fig 1. It is conceivable that to have better quality images, a pre-processing method was employed. After obtaining new images, a feature extractor was applied to have better data quality and the best and most important features, a feature selector method, namely, Lasso, was employed. Finally, an ensemble method was utilized to increase evaluation metrics.

Table 1. Detailed summaries on predicting and classifying skin cancer in other studies.

References	Dataset	Method
Kumar K et al. [5]	Kaggle+ CIA datasets	AuDNN+ IOT
Balaha and Hassan [6]	ISIC (2016-2017-2018)	Transfer learning and sparrow search algorithm
Tabrizchi et al. [7]	SIIM-ISIC	VGG
Adla et al. [8]	ISIC 2019+ ISIC 2020	FrCN-DGCA
Mridha et al. [9]	HAM10000	Optimized Convolutional Neural Network
Wu et al. [10]	HAM10000	Standard convolution+ wavelet down-sampling
Qasim Gilani et al. [11]	ISIC 2019	Deep Spiking Neural Network
Huang et al. [12]	ISIC Library	YOLOv5
Teodoro et al. [13]	ISDIS	GAN+ RoI-Based
S M et al. [14]	ISIC-2020	DCNN
Mawgoud et al. [15]	MODE-NODE, ISIC	Deep Transfer Learning
Khan et al. [16]	ISBI2016, ISIC2017, HAM10000	Deep learning
Khan et al. [17]	ISBI2016, ISBI2017	Deep Learning with IcNR-based feature selection
Khan et al. [18]	PH2, ISBI2016, ISIC2017, HAM10000	Mask R-CNN, CNN

<https://doi.org/10.1371/journal.pone.0301275.t001>

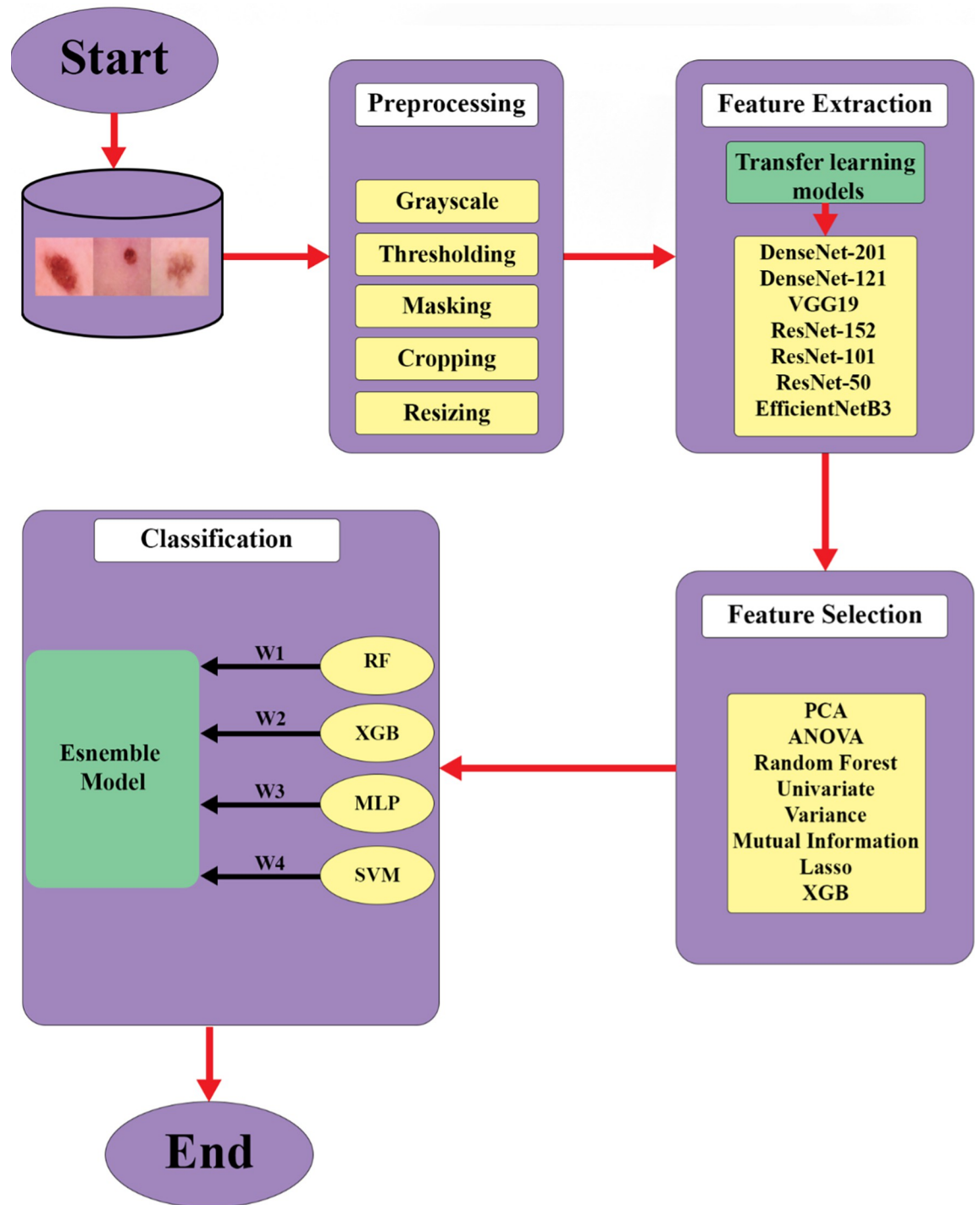


Fig 1. The workflow of the proposed method.

<https://doi.org/10.1371/journal.pone.0301275.g001>

The novelty in this paper lies in the comprehensive and systematic exploration and combination of these various methodologies, suggesting an in-depth approach to feature extraction, feature selection, and model combination to boost the overall performance.

This research uses a dataset [19] that has a harmoniously balanced compilation of images portraying both benign and malignant skin moles. The dataset is structured into two distinct folders, all captured at a resolution of 224x224 pixels. These images vividly depict the contrasting characteristics of the two-mole types, offering a comprehensive visual resource for discerning the nuances between benign and malignant skin anomalies.

Preprocessing

The preprocessing of dermoscopic images is paramount in developing precise algorithms for automated skin cancer diagnosis [20]. Skin cancer is among the most widespread types of cancer globally [21]. Early and accurate diagnosis is vital for improving prognoses and survival rates. However, the accuracy of diagnosis heavily relies on the expertise of dermatologists who visually assess and analyze skin lesions from dermoscopic images. Automated algorithms seek to enhance diagnostic precision by utilizing computerized analysis of skin images. Nevertheless, original dermoscopic images often contain flaws such as noise, inconsistent lighting, and other irregularities that can adversely affect algorithm performance. Proficient preprocessing of these images is instrumental in standardizing the data and bolstering algorithm dependability.

The primary objective of preprocessing in skin cancer diagnosis is to enhance image quality and prepare the data for in-depth analysis by the diagnostic algorithm. Fundamental preprocessing steps encompass the conversion of color spaces, removal of hair, correction of uneven illumination, enhancement of contrast, and elimination of noise. The shift from RGB color space to alternative color spaces like CIELAB aids in normalizing skin color and lesion patterns [22]. Noises caused by hair-over lesions are addressed through inpainting, isolating the lesions themselves. Uneven illumination and reflections that may obscure lesions are mitigated through illumination correction, ensuring uniform brightness and heightened contrast. Filtering the image is also pivotal for eradicating extraneous artifacts and balancing noise. Employing these preprocessing techniques helps isolate diagnostically significant features of the lesion for the algorithm to analyze.

Following the fundamental preprocessing steps, additional techniques contribute to standardizing and normalizing specific features within the prepared images. Strategies such as histogram equalization bolster contrast and amplify feature visibility. Normalizing color, size, and structure in the preprocessed images empowers the algorithm to characterize, categorize, and compare lesions more accurately. Data augmentation through rotation, scaling, and flipping can amplify dataset size and variability, enhancing algorithm training. Apt preprocessing produces refined, normalized image data, aiding complex algorithms in extracting subtle features and patterns for precise diagnosis.

In essence, meticulous preprocessing of dermoscopic images diminishes noise, and distracting variability while accentuating the visually meaningful diagnostic features of skin lesions. Standardized, improved images facilitate machine learning algorithms in effortlessly identifying patterns indicative of cancerous or benign lesions. By establishing robust preprocessing pipelines tailored to the specific challenges of dermoscopic images, researchers can formulate highly accurate algorithms for automated skin cancer screening and diagnosis. Dependable preprocessing facilitates the seamless real-world integration of these systems, augmenting public access, efficiency, and the quality of skin cancer diagnosis.

In this research, RGB-formatted images transform grayscale. Subsequently, a binary inverse thresholding operation using Otsu's method is applied. This operation provides a binary mask that accentuates the object of interest against its backdrop [23]. Employing this mask, the object is then segregated from the background within the RGB image through bitwise

operations. Pixels representing the background are altered to white, erasing the background itself. Through the identification of indices corresponding to non-zero (foreground) pixels in the processed image, the parameters of the bounding box enclosing the object are discerned. This information facilitates the extraction of a cropped variant of the RGB image, encompassing solely the object. The cropped image is subsequently resized to conform to a standardized dimension of 224x224 pixels, ensuring uniformity in input measurements. Ultimately, the resized image is incorporated into a roster of images designated for subsequent analysis. This meticulous preprocessing sequence culminates in the creation of a collection of images, uniformly sized and centered. These images are poised for subsequent computational tasks, such as image classification or object recognition.

Feature extraction

Extracting features is a crucial stage in advancing machine learning systems for automated skin cancer diagnosis. While preprocessing of dermoscopic images normalizes and enriches the initial data, feature extraction precisely identifies and segregates the distinct visual attributes and patterns present in the lesions. Thoughtful selection and extraction of only the most diagnostically pertinent features facilitate more effective and focused analysis, concurrently reducing complexity for the classification algorithm. The objective is to derive a set of features that can faithfully represent each image, enabling the system to discern discerning patterns distinguishing between malignant melanomas and harmless lesions.

Features at the pixel level, such as color, texture, and shape, alongside more complex semantic features describing asymmetry, border irregularity, color variations, and dimensions, offer potential value in distinguishing cancerous and non-cancerous skin lesions. The choice of extraction techniques depends on the priority of features in the computer vision approach. For pixel-level features, standard methods involve segmentation, thresholding, edge detection, and filtering to isolate colors, textures, and structures. Segmentation divides the image into regions with comparable traits [24]. Thresholding emphasizes pixel intensities indicative of boundaries and shapes [25]. Edge detection identifies lines and abrupt changes in the image [26]. Gabor filters and other filtering methods characterize texture patterns [27].

In contrast, extracting higher-level diagnostic features necessitates algorithms to quantify semantic attributes and patterns in the lesions. These algorithms can pinpoint and measure asymmetry across multiple axes of the lesion. Algorithms gauging border irregularity evaluate the smoothness versus jaggedness of the lesion edges. Others assessing color variations can classify multiple colors and uniformity of color across the lesion area. Algorithms for dimensions extract size characteristics and dimensional ratios. A meticulous selection of a concise set of the most valuable features for diagnosis prevents unnecessary complexity in the model while enabling it to discern the visual patterns that most consistently set apart malignant and benign skin lesions.

Strategic feature extraction empowers machine learning models to pinpoint and concentrate on the most diagnostically meaningful elements within dermoscopic images. Condensing each image to its essential array of distinguishing visual features augments efficiency and precision while minimizing model intricacies. Coupled with robust preprocessing, deliberate feature extraction provides optimized data for algorithm training and evaluation in automated skin cancer diagnosis. The use of more focused data enhances model performance and facilitates integration into actual clinical settings, thus advancing dermatological care.

In the realm of skin cancer analysis, the integration of pre-trained convolutional neural network (CNN) models as feature extractors has ushered in a transformative approach to enhancing diagnostic accuracy. Amidst the diverse array of models employed for this purpose, such as

DenseNet-121 [28], VGG19 [29], ResNet-50, ResNet-101, ResNet-152 [30], and EfficientNet-B3 [31], a rigorous performance evaluation process spotlighted DenseNet-201 as the preeminent contender. These pre-trained CNN models, having undergone comprehensive training on extensive image datasets, exhibit an innate capability to learn and encapsulate hierarchical features from input images autonomously. This intrinsic trait enables them to undertake efficient feature extraction sans the need for labor-intensive manual feature engineering. Of noteworthy significance, DenseNet architectures, characterized by densely interwoven layers, proffer an exclusive advantage by fostering unimpeded information flow among layers bolstered by intricate connectivity patterns. This unique architectural structure equips the models with an exceptional ability to simultaneously capture both intricate details at the lower level and complicated, complex features at the higher level. As a result, these models are highly proficient in detecting subtle nuances present within dermatological images—nuances that often hold critical significance as indicators of potential skin cancer.

Feature selection

While feature extraction detects promising attributes within images, feature selection takes a step further by narrowing down the data to the most pertinent characteristics crucial for skin cancer diagnosis. Not all the extracted features have valuable insights to distinguish between malignant melanomas and harmless lesions. Feature selection assesses the extracted features and cherry-picks a subset that possesses the optimal discriminating potential for precise classification. This streamlining process enhances the efficiency, accuracy, and comprehensibility of the model by eliminating superfluous, inconsequential, or bewildering features from the dataset.

Efficient feature selection mandates analytical techniques to gauge and rank the utility of the extracted features. Commonly, methods like filters, wrappers, and embedded approaches are employed for this purpose. Filter techniques use statistical metrics like correlation, mutual information, or chi-square tests to score and rank features independently of the model [32]. Wrappers utilize the model itself to test various feature combinations and assess the model's performance [33]. Embedded methods conduct feature selection as an integral part of the model construction process [34]. In the context of skin cancer diagnosis, a blend of these techniques is often necessary to pinpoint the most compelling features.

Prudent feature selection scrutinizes texture, color, shape, and semantic attributes to ascertain which offers the highest discriminatory potential. The optimal colors distinguishing melanomas from benign lesions are singled out. Specific texture patterns can indicate malignancy. Asymmetry, border irregularity, and dimensional ratios with notable discriminatory capabilities are retained, while redundant features are discarded. This streamlining and reduction of the feature space facilitate more efficient algorithm training, preventing distraction from irrelevant inputs.

By removing redundant and non-informative inputs, feature selection guides the model to concentrate solely on the most relevant components essential for optimal performance. This not only enhances model interpretability but also augments predictive accuracy. The outcome is an optimized subset of features that allows effective training of machine learning models for automated analysis of dermoscopic images, ultimately improving melanoma detection.

Following the phase of extracting features using DenseNet-201 in this study, a crucial step is introduced to enhance and optimize the obtained feature set. In this context, a significant technique known as the Lasso method is introduced for feature selection. Lasso, or Least Absolute Shrinkage and Selection Operator, holds prominence as a sophisticated regularization method that achieves both feature selection and coefficient reduction. Operating within the framework of linear regression, Lasso integrates a penalty term that encourages specific coefficients to be precisely reduced to zero, effectively removing the associated features [35]. The

utilization of Lasso in this investigation involves a thorough assessment of the relevance of each extracted feature concerning the classification task. By capitalizing on Lasso's inherent ability to effectively trim irrelevant or redundant features while retaining the most informative ones, the feature set is refined to encompass only those features that substantially contribute to the given task. This refinement enhances the model's generalization capability and reduces overfitting risks. The integration of Lasso as a feature selector enhances the overall efficiency and comprehensibility of the classification model. Through this unified approach, the study aims to attain an improved feature representation that not only captures the complexities of dermatological images but also leads to enhanced performance in skin cancer classification.

Classification

The accurate categorization of skin lesions as either malignant or benign stands as a pivotal aspect in the creation of dependable automated frameworks for skin cancer screening and diagnosis [36]. While ensuring the quality of image pre-processing, meticulous feature extraction and judicious filtering of features help optimize the input data, the onus of making the ultimate diagnostic determination lies with the classification algorithm. Given the global prevalence of skin cancer, including melanoma, enhancing screening and diagnostic precision holds tremendous advantages concerning early detection, treatment, and survival rates. Advanced classification algorithms possess the potential to either equal or potentially surpass the diagnostic capabilities of human dermatologists by discerning subtle patterns within dermoscopic images.

The classification algorithm thoroughly evaluates the refined set of visual features extracted from the pre-processed dermoscopic images and sets decision boundaries to classify new samples as either malignant or benign consistently. A broad array of machine learning classification methodologies can be experimented with and fine-tuned to pinpoint the most effective diagnostic approach for the skin lesion classification task. Commonly employed classification algorithms encompass logistic regression, support vector machines, random forests, artificial neural networks, gradient boosting machines, and diverse ensembles. Logistic regression models gauge probability scores based on the feature inputs [37], indicating the likelihood of malignancy or benignity. Support vector machines ascertain optimal hyperplanes between classes in the multivariate feature space [38]. Random forests construct collections of decision trees based on random subsets of features to amplify the overall performance [39]. Artificial neural networks utilize hidden layers to capture intricate nonlinear relationships between image features and diagnostic outcomes.

Within each algorithm, exhaustive parameter adjustment through grid or random searches identifies the optimal model settings and architecture to maximize classification accuracy. Fine-tuning kernels, regularization terms, tree depths, hidden layers, activation functions, and other parameters offer flexibility to tailor the model's intricacy to suit the challenge posed by the skin image data. Robust scoring metrics like AUC-ROC, precision, recall, and F1-score through cross-validation prevent overfitting and provide an objective understanding of model generalizability. The most effective classification model can then be solidified by training on the entire refined dataset.

After being rigorously optimized through this stringent process, the classification model can consistently categorize new dermoscopic images with a high degree of accuracy. This paves the way for the development of fully automated skin cancer screening systems that can be clinically validated before responsible real-world implementation. These systems have demonstrated the potential to enhance diagnostic speed, objectivity, accessibility, and accuracy. Prudent and responsible use of AI-based classification has the potential to significantly benefit skin cancer patients through early detection, cost reduction, improved outcomes, and

potentially saving lives. Although challenges remain for full integration into clinical practice, advanced machine learning classification algorithms hold immense promise to assist dermatologists and enhance skin cancer diagnosis.

Ensemble techniques, a prominent facet within the realm of machine learning, have garnered considerable attention for their exceptional capacity to enhance predictive accuracy and broaden applicability across diverse domains, including the nuanced domain of medical image analysis. The underlying principle of these methods involves amalgamating multiple individual models to craft a more resilient and precise ensemble model. Particularly in the realm of skin cancer detection, where accuracy and precision hold paramount importance, ensemble methods offer an efficacious avenue to elevate the performance of machine learning algorithms. A range of ensemble methodologies exists, each distinguished by unique attributes contributing to their efficacy. For instance, Bagging (Bootstrap Aggregating), demonstrated through the example of Random Forest, entails constructing numerous decision tree models by training them on distinct subsets of the training data [40]. These models' predictions are subsequently consolidated to have a definitive outcome, effectively countering overfitting while augmenting robustness. On the other hand, Boosting iteratively trains modest models, such as shallow decision trees, by assigning elevated weights to misclassified instances [41]. This iterative approach culminates in a robust ensemble model. Stacking, a more intricate approach, encompasses training multiple foundational models, subsequently training a meta-model to proficiently amalgamate the predictions of these foundational models [42]. The Random Subspace Method (Feature Bagging) entails training each foundational model on random feature subsets, a strategy that effectively mitigates the risk of overfitting [43]. In the context of skin cancer detection, ensemble techniques hold considerable promise, primarily attributable to the intricacies and diverseness inherent in dermatological images. The task of distinguishing between benign and malignant lesions necessitates capturing intricate patterns that a solitary algorithm might struggle to encapsulate fully. Here, ensemble methods emerge as a viable solution by amalgamating the strengths of multiple algorithms, thus countering individual shortcomings and augmenting overall accuracy. In the research endeavor, an ensemble comprising diverse machine-learning algorithms was harnessed for skin cancer detection. This ensemble encompassed XGBoost, renowned for its gradient-boosting mechanism that iteratively hones the model's predictions, giving precedence to misclassified instances [44]. MLP (Multi-Layer Perceptron), a neural network architecture with multiple concealed layers capable of capturing intricate data relationships, contributed to its distinctive prowess [45]. The MLP architecture utilized in this study embodies a feedforward neural network structure, comprising distinct layers tailored for the effective classification of skin cancer lesions. The input layer encapsulates the flattened representation of extracted image features, aligning with the dimensions of the resized and preprocessed skin lesion images. Throughout the architecture, multiple hidden layers, characterized by varying neuron counts, facilitate the learning of intricate patterns inherent in the input data. Employing ReLU activation functions within these hidden layers ensures efficient gradient propagation and fosters rapid convergence during model training. Additionally, the output layer is equipped with a sigmoid activation function, facilitating the generation of probabilistic outputs suitable for binary classification tasks. Alongside these core components, the MLP architecture integrates dropout layers to mitigate overfitting, batch normalization techniques to stabilize training dynamics, and regularization mechanisms to enhance model generalization. These architectural nuances collectively empower the MLP to discern nuanced features and achieve robust classification performance, thus rendering it a pivotal asset in the realm of skin cancer detection through medical imaging. The ensemble further incorporated SVM (Support Vector Machine), celebrated for its proficiency in segregating data [46] into distinct classes by defining a hyperplane, and Random Forest, an assemblage of decision trees

acclaimed for their robust generalization capacity [39]. Each of these algorithms contributed unique strengths to the ensemble's arsenal. The ensemble amalgamation of XGBoost, MLP, SVM, and Random Forest within this research exemplifies the potential of such a collaborative approach. By synergistically harnessing the individual strengths of these algorithms, the ensemble aimed to attain elevated precision and robustness in detecting malignant lesions. This collaborative strategy is particularly powerful in incorporating a diverse array of features, thereby augmenting the model's efficacy in discerning between benign and malignant cases. In a rapidly evolving landscape of machine learning, ensemble methods stand as a steadfast route to expanding the horizons of medical image analysis and elevating patient care.

Discussion and results

In this section, a detailed analysis of the middle steps, process, and performance of the proposed technique for skin cancer detection using machine learning and deep learning methods is presented. Beginning with an examination of each stage's intricacies, followed by a comprehensive evaluation of the performance metrics and a comparison with alternative methods.

The outcomes of employing various pre-trained transfer learning models and distinct classification algorithms were elucidated. The techniques outlined in this paper are reproduced in a Google Colab, making use of the Python programming language.

The proposed technique encompasses several key stages, each critical for achieving accurate and reliable skin cancer detection. Image preprocessing involves advanced techniques to enhance image quality and remove noise, addressing common challenges such as illumination variations and artifacts inherent in dermatological images. Specifically, Otsu's method was employed to ensure high-quality data for subsequent processing.

Feature extraction utilized pre-trained deep learning models such as DenseNet-201 to capture discriminative features from dermoscopic images. Additionally, the Lasso method was employed as a feature selector to streamline the feature space and maximize the discriminative potential of the extracted features.

The classification stage involved the application of diverse algorithms, including XGBoost, Multi-Layer Perceptron (MLP), Support Vector Machine (SVM), and Random Forest. Each algorithm was fine-tuned and optimized to delineate between malignant and benign skin lesions, with careful consideration given to parameter tuning and model architecture.

Rigorous evaluation of the proposed technique assessed its performance in skin cancer detection using standard evaluation metrics such as accuracy, precision, recall, and F1-score. A comparative analysis with alternative methods showcased the superior performance of the technique relative to existing approaches. DenseNet-201 emerged as the optimal feature extractor, outperforming other pre-trained models such as ResNet and VGG19.

The evaluation of this research is predicated upon its classification performance, gauged by different evaluation metrics. This study uses accuracy, precision, recall, and F1-score to assess the performance of the proposed method. All the performance metrics are computed as shown in [Table 2](#).

Table 2. The performance metrics.

Parameter	Value
Accuracy	$\frac{\text{True Positive} + \text{True Negative}}{\text{True Positive} + \text{True Negative} + \text{False Positive} + \text{False Negative}}$
Precision	$\frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$
Recall	$\frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$
F1-Score	$2 \cdot \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$

<https://doi.org/10.1371/journal.pone.0301275.t002>

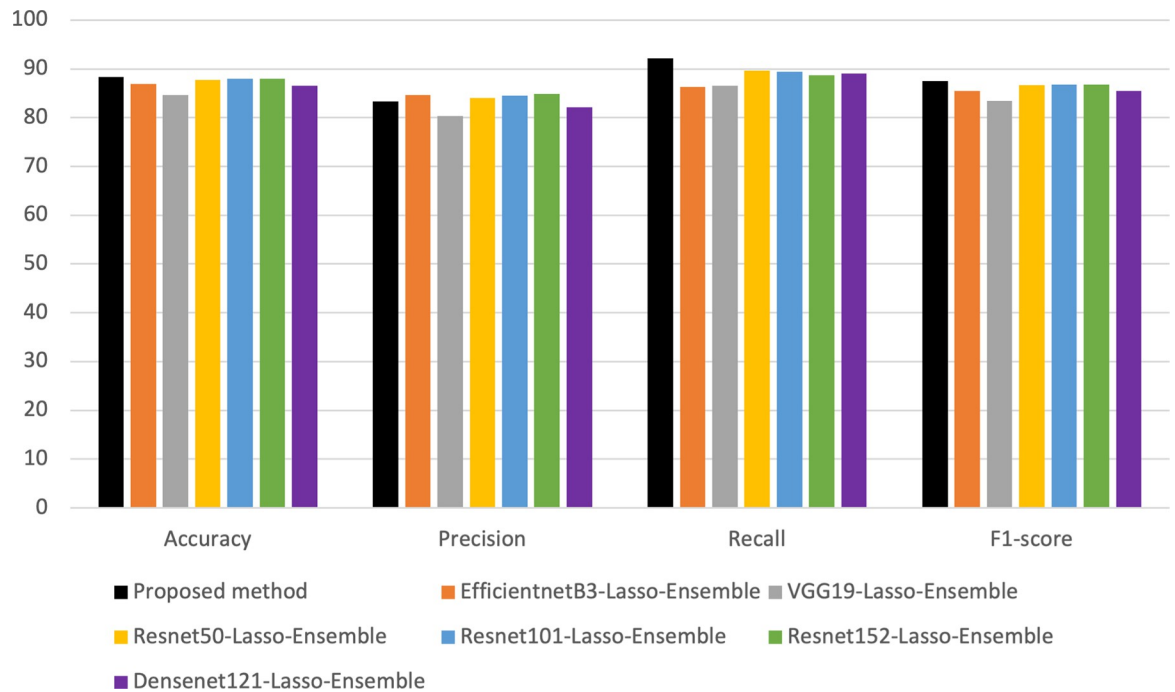


Fig 2. Comparison of different feature extractors.

<https://doi.org/10.1371/journal.pone.0301275.g002>

Following a comprehensive comparative analysis, the superiority of DenseNet-201 consistently emerged. The model's capability to comprehend and encapsulate finely detailed attributes played a pivotal role in significantly enhancing precision when distinguishing between malignant and benign skin lesions. It is this exceptional aptitude for capturing intricate aspects that firmly establish the prominence of DenseNet-201 as the optimal choice among its counterparts. Fig 2. illustrates a comparative evaluation of DenseNet-201 concerning other pre-trained models employed in the study.

Based on the observations derived from Fig 2., it can be deduced that the outcomes obtained through the proposed methodology exhibit superior performance when compared to alternative techniques, specifically the Resnet family, EfficientnetB3, and VGG19.

After finding DenseNet-201 as the best feature extractor, Lasso, as a feature selector, provides a streamlined and optimized subset of features that maximizes the distinguishing potential of the extracted features. In this study, an extensive ensemble strategy was implemented to enhance the accuracy of skin cancer detection. The ensemble involved the integration of various machine learning models, encompassing XGBoost (XGB), Multi-Layer Perceptron (MLP), Support Vector Machine (SVM), and Random Forest. The stacking technique was applied for the ensemble process, merging predictions from individual models to provide a conclusive decision.

Fig 3. provides a visual representation of the metrics' improvement resulting from implementing the proposed methodology.

The ensemble strategy, incorporating diverse machine learning models, further enhanced the accuracy and robustness of the approach. Fig 3. reveals that the proposed method excels particularly in the recall metric, which is among the most important in the healthcare context [47]. More specifically, the achieved recall rate of 92.15% surpasses that of alternative methodologies, while the attained accuracy rate of 87.72% also demonstrates superior performance relative to competing approaches.

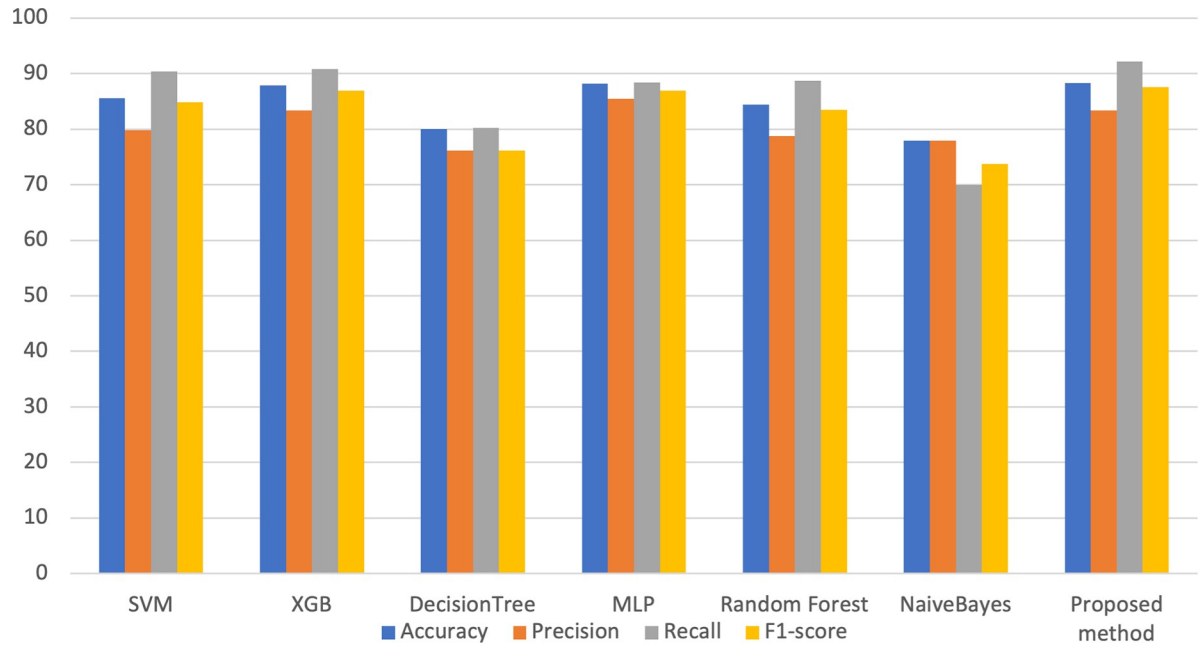


Fig 3. Obtained metrics in the proposed method and other applied classifiers.

<https://doi.org/10.1371/journal.pone.0301275.g003>

Two types of ensembles, whose results are shown in Fig 4., were applied, in which the proposed method was chosen because of its potential.

Fig 4. is a comprehensive comparative analysis to assess the performance outcomes by the novel methodology in contrast to a basic ensemble approach. Findings indicate that the proposed method outperforms the simple ensemble method in three out of four evaluation metrics.

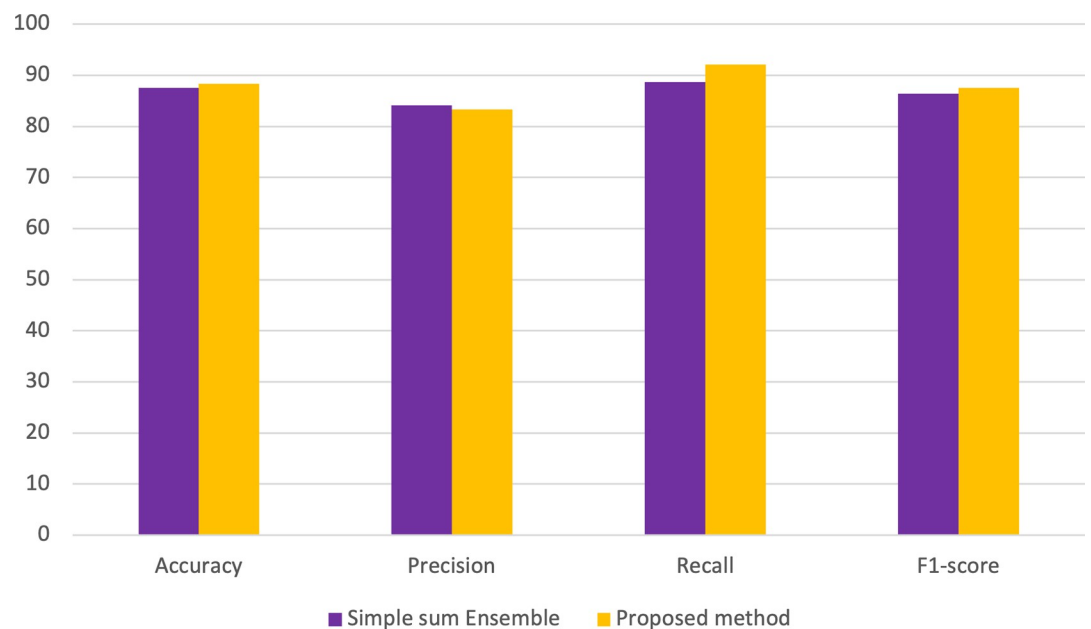


Fig 4. A comparative analysis of the achieved outcomes derived from the proposed methodology in comparison with a straightforward ensemble approach.

<https://doi.org/10.1371/journal.pone.0301275.g004>

Notably, the importance of each model was acknowledged through the allocation of weights proportionate to their respective accuracies. Notably, the MLP model, exhibiting the highest accuracy within the ensemble, received the greatest weight. Subsequently, XGBoost, while offering accuracy slightly below MLP but surpassing SVM and Random Forest, was assigned the second-highest weight. SVM, ranking third in terms of accuracy, followed by Random Forest, with the fourth-highest accuracy, received progressively reduced weights within the ensemble framework. The stacking ensemble technique effectively harnesses the collective capabilities of diverse machine learning models. The evaluation metrics acquired are depicted in Fig 5.

As illustrated in Fig 5., the metrics obtained in the study demonstrate promising applicability within the realm of healthcare. By strategically determining weights based on accuracy, the stacking approach optimally exploits the distinct strengths of individual models, enhancing the final decision-making process. This ensemble method emphasizes a tactical fusion of models, wherein their cooperative outcomes surpass the performance of unique models. This convergence results in a more robust and refined diagnostic system, amplifying its effectiveness in detecting skin cancer.

A confusion matrix is a performance measurement tool for classification problems, displaying the number of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) predictions made by a classification model. It provides a comprehensive overview of a model's accuracy, precision, recall, and F1-score, facilitating performance evaluation, model comparison, error analysis, threshold selection, and class imbalance handling. By analyzing TP, TN, FP, and FN rates, identify common prediction errors, optimize decision thresholds, and mitigate bias in datasets with class imbalance, enhancing the credibility of research findings in the field of machine learning. In Fig 6. the confusion matrix achieved in this paper is visualized.

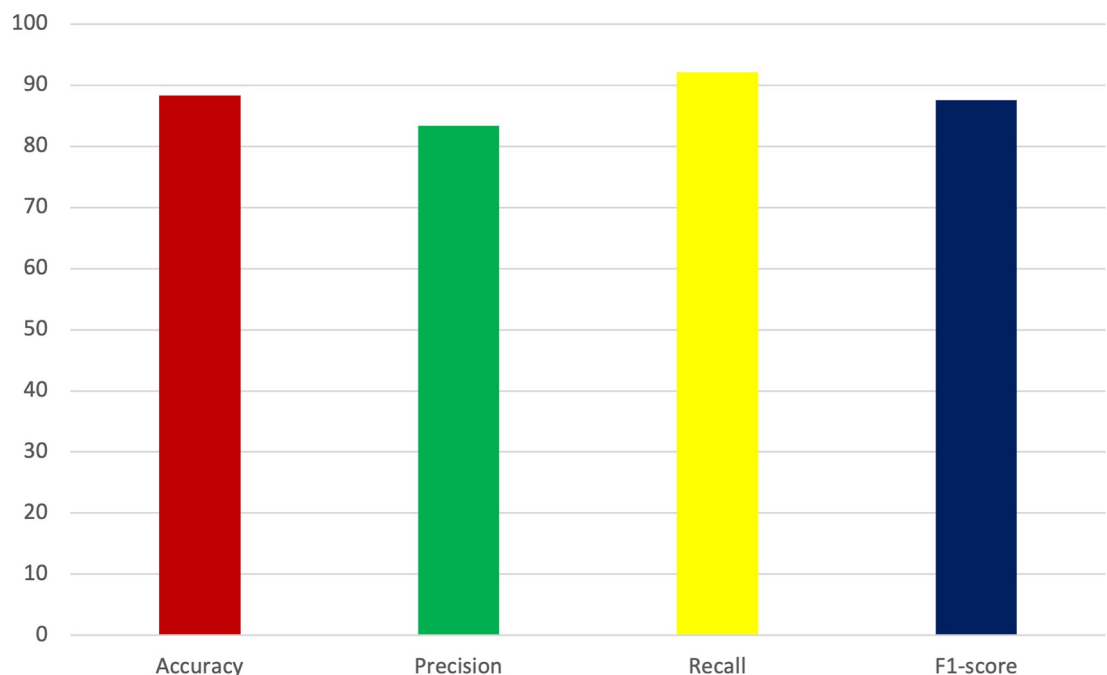


Fig 5. Obtained metrics in the proposed method and other applied classifiers.

<https://doi.org/10.1371/journal.pone.0301275.g005>

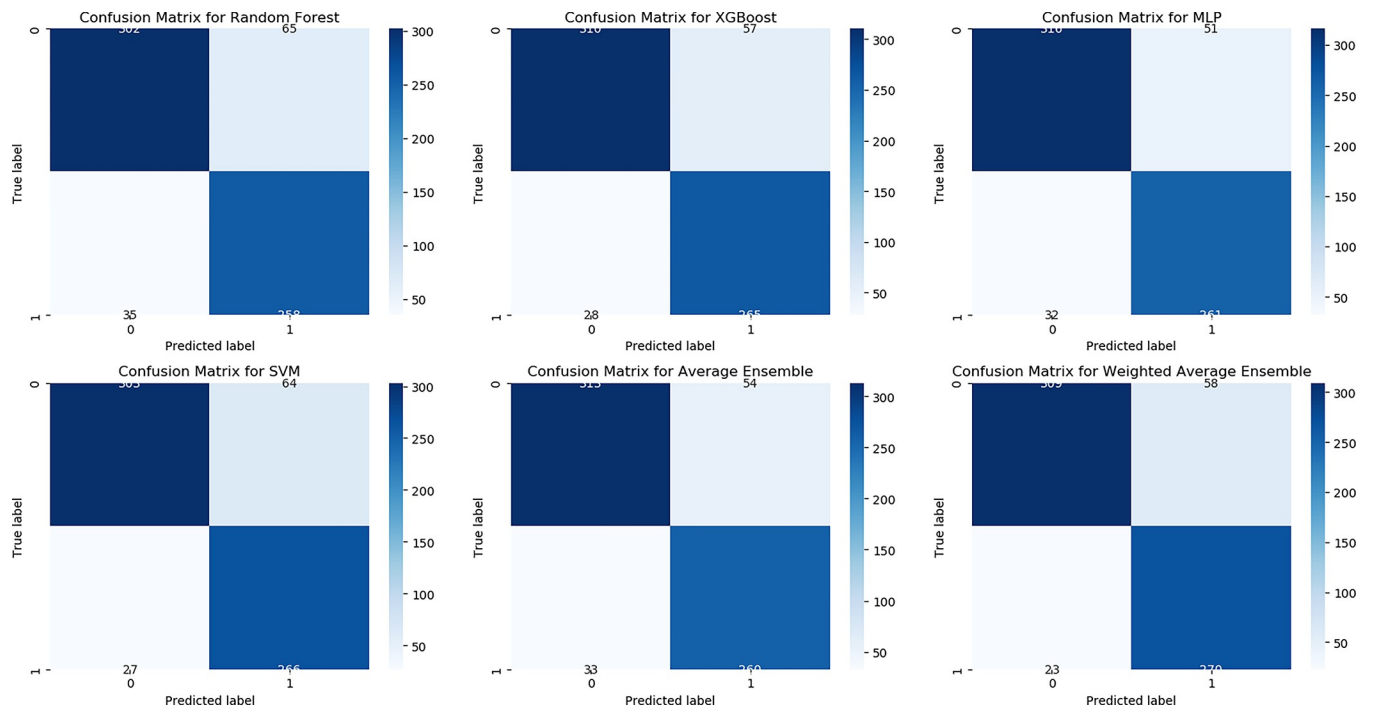


Fig 6. Confusion matrix in used classifiers.

<https://doi.org/10.1371/journal.pone.0301275.g006>

The ensemble methodology adopted in this study holds significant potential for advancing medical image analysis, particularly within the domain of dermatological imagery. By judiciously amalgamating models with varying strengths and employing stacking as the mechanism of amalgamation, this research seeks to elevate the precision of skin cancer detection. This endeavor, in turn, augments the potential for accurate and timely medical intervention, reflecting a promising stride toward improved patient care.

Comparison with cutting-edge models

Table 3 presents a comparison of diverse studies employing different methods to classify skin cancer along with their respective outcomes. Mijwil [48] utilized InceptionV3 CNN to classify skin cancer images with 86.90% accuracy, demonstrating its effectiveness in distinguishing benign and malignant cases. Jasil and Ulagamuthalvi [49] explored pre-trained models like VGG16, VGG19, and Inception V3, achieving 74–77% accuracy in classifying skin lesions and highlighting their potential in medical image processing. Dubal et al. [50] introduced a technique using ordinary cameras and Neural Networks to accurately identify cancerous skin abnormalities through visual segmentation. Brinker et al. [51] focused on the binary categorization of skin abnormalities using dermoscopic images and annotations from specialists. Majtner et al. [52] developed an improved melanoma detection method combining deep learning with LDA for feature reduction and enhanced classification accuracy. Gupta et al. [53] presented a methodology enhancing dynamic training and testing procedures for more pronounced outcomes, emphasizing resource-intensive augmentation strategies. Murugan et al. [54] addressed skin cancer diagnosis, focusing on melanoma, employing watershed segmentation and various classifiers, with SVM proving the most effective. Mustafa et al. [55] introduced a method using color space, luminance, and SVM for categorizing benign/malignant lesions based on shape characteristics, emphasizing the SVM's capability. Linsangan and

Table 3. Comparison with other studies.

References	Dataset	Method and Methods Used	Accuracy	Sensitivity
Jasil and Ulagamuthalvi [49]	ISIC	Transfer learning	77	-
Dubal et al. [50]	-	Neural Networks	76.9	-
Brinker et al. [51]	ISBI2016	CNN	83.9	56
Gupta et al. [53]	ISIC2016	Ensemble techniques	87	67
Murugan et al. [54]	ISIC	Random Forest	76.87	78.43
Mustafa et al. [55]	-	SVM	80.00	71.43
Linsangan and Adtoon [56]	ISIC archive	KNN	86.67	-
Shalu and Kamboj [57]	MED-NODE	Decision Tree	82.35	-
Proposed method	ISIC	DenseNet-201+ Lasso+ Ensemble learning	87.72	92.15

<https://doi.org/10.1371/journal.pone.0301275.t003>

Adtoon [56] focused on non-invasive skin cancer detection, achieving 86.67% accuracy using geometric features on a Raspberry Pi. Shalu and Kamboj [57] developed a melanoma detection system using preprocessing and segmentation techniques. With an accuracy of 87.72% and sensitivity of 92.15%, the results of this paper surpass many of the mentioned methodologies, signifying a notable enhancement in the accuracy and sensitivity of skin cancer based on the employed techniques.

Conclusion

This paper aims to assist medical professionals in accurately distinguishing between benign and malignant skin cancer cases using machine learning and deep learning methods. Among the pre-trained models, including ResNet-50, ResNet-101, ResNet-152, VGG19, EfficientNet-B3, and DenseNet-121, DenseNet-201 has been chosen as the feature extractor. In the subsequent phase, the Lasso method is employed as a feature selector, followed by the implementation of diverse machine learning techniques. The innovation lies within the meticulous and systematic investigation and fusion of diverse methodologies, signifying a profound approach towards extracting features, selecting crucial elements, and amalgamating models. The best accuracy and precision are selected and improved through ensemble methods. Given the importance of the sensitivity metric in disease detection, efforts have been directed toward achieving a substantial value. The accuracy and sensitivity values achieved in this study are 87.72% and 92.15%, respectively. The clinical relevance and potential impact of the proposed technique on dermatological practice are profound. By leveraging state-of-the-art machine learning and deep learning methods, a reliable and efficient solution for early detection and diagnosis of skin cancer is offered. The approach holds promise for reducing healthcare costs. In future research endeavors, further refinement and validation of the technique using diverse datasets and advanced deep-learning architectures are planned. Additionally, exploration of the integration of complementary modalities such as clinical data and histopathological findings to enhance diagnostic accuracy is underway.

Acknowledgments

This study is supported via funding from Prince Satam bin Abdulaziz University project number (PSAU/2024/R/1445)

Author Contributions

Conceptualization: Firas Muhammad Zeki Mahmood, Aso Darwesh.

Data curation: Dildar Hussain, Farhan A. Alenizi, Amirhossein Noroozi Varzeghani, Parvaneh Asghari.

Methodology: Mehdi Hosseinzadeh, Mazhar Hussain Malik.

Project administration: Sang-Woong Lee.

Writing – original draft: Mehdi Hosseinzadeh, Dildar Hussain, Firas Muhammad Zeki Mahmood, Farhan A. Alenizi, Amirhossein Noroozi Varzeghani, Parvaneh Asghari, Aso Darwesh, Mazhar Hussain Malik, Sang-Woong Lee.

References

1. Panieri E, Telkoparan-Akillilar P, Saso L. NRF2, a crucial modulator of skin cells protection against vitiligo, psoriasis, and cancer. *BioFactors*. 2023; 49(2):228–50. <https://doi.org/10.1002/biof.1912> PMID: 36310374
2. Sutopo NC, Kim JH, Cho JY. Role of histone methylation in skin cancers: Histone methylation-modifying enzymes as a new class of targets for skin cancer treatment. *Biochim Biophys Acta BBA—Rev Cancer*. 2023 May; 1878(3):188865. <https://doi.org/10.1016/j.bbcan.2023.188865> PMID: 36841366
3. Algarni AM, Alshehri HS, Al Zomia AS, Alhifithi MA, Lahiq LA, Al Fae FM, et al. The Epidemiological Pattern of Skin Cancer from 2011 to 2022 among the Population of the Aseer Region, Kingdom of Saudi Arabia. *Cancers*. 2023 Jan; 15(18):4612. <https://doi.org/10.3390/cancers15184612> PMID: 37760581
4. Esteva A, Kuprel B, Thrun S. Deep networks for early stage skin disease and skin cancer classification. *Proj Rep Stanf Univ*. 2015;
5. Kumar K A, Satheesha TY, Salvador BBL, Mithileysh S, Ahmed ST. Augmented Intelligence enabled Deep Neural Networking (AuDNN) framework for skin cancer classification and prediction using multi-dimensional datasets on industrial IoT standards. *Microprocess Microsyst*. 2023 Mar 1; 97:104755.
6. Balaha HM, Hassan AES. Skin cancer diagnosis based on deep transfer learning and sparrow search algorithm. *Neural Comput Appl*. 2023 Jan; 35(1):815–53.
7. Tabrizchi H, Parvizpour S, Razmara J. An Improved VGG Model for Skin Cancer Detection. *Neural Process Lett*. 2023 Aug 1; 55(4):3715–32.
8. Adla D, Reddy GVR, Nayak P, Karuna G. A full-resolution convolutional network with a dynamic graph cut algorithm for skin cancer classification and detection. *Healthc Anal*. 2023 Nov 1; 3:100154.
9. Mridha K, Uddin MdM, Shin J, Khadka S, Mridha MF. An Interpretable Skin Cancer Classification Using Optimized Convolutional Neural Network for a Smart Healthcare System. *IEEE Access*. 2023; 11:41003–18.
10. Wu Q e, Yu Y, Zhang X. A Skin Cancer Classification Method Based on Discrete Wavelet Down-Sampling Feature Reconstruction. *Electronics*. 2023 Jan; 12(9):2103.
11. Qasim Gilani S, Syed T, Umair M, Marques O. Skin Cancer Classification Using Deep Spiking Neural Network. *J Digit Imaging*. 2023 Jun 1; 36(3):1137–47. <https://doi.org/10.1007/s10278-023-00776-2> PMID: 36690775
12. Huang HY, Hsiao YP, Mukundan A, Tsao YM, Chang WY, Wang HC. Classification of Skin Cancer Using Novel Hyperspectral Imaging Engineering via YOLOv5. *J Clin Med*. 2023 Jan; 12(3):1134. <https://doi.org/10.3390/jcm12031134> PMID: 36769781
13. Teodoro AAM, Silva DH, Rosa RL, Saadi M, Wuttisittikulij L, Mumtaz RA, et al. A Skin Cancer Classification Approach using GAN and RoI-Based Attention Mechanism. *J Signal Process Syst*. 2023 Mar 1; 95(2):211–24.
14. S M J P M, Aravindan C, Appavu R. Classification of skin cancer from dermoscopic images using deep neural network architectures. *Multimed Tools Appl*. 2023 Apr; 82(10):15763–78. <https://doi.org/10.1007/s11042-022-13847-3> PMID: 36250184
15. Mawgoud AA. Empirical Study and Enhancement on Deep Transfer Learning for Skin Lesions Detection. 2020 Nov 21 [cited 2024 Mar 7]; Available from: https://www.academia.edu/79955996/Empirical_Study_and_Enhancement_on_Deep_Transfer_Learning_for_Skin_Lesions_Detection
16. Khan MA, Akram T, Zhang YD, Sharif M. Attributes based skin lesion detection and recognition: A mask RCNN and transfer learning-based deep learning framework. *Pattern Recognit Lett*. 2021 Mar 1; 143:58–66.
17. Khan MA, Sharif M, Akram T, Bukhari SAC, Nayak RS. Developed Newton-Raphson based deep features selection framework for skin lesion recognition. *Pattern Recognit Lett*. 2020 Jan 1; 129:293–303.

18. Khan MA, Zhang YD, Sharif M, Akram T. Pixels to Classes: Intelligent Learning Framework for Multi-class Skin Lesion Localization and Classification. *Comput Electr Eng*. 2021 Mar 1; 90:106956.
19. ISIC [Internet]. [cited 2023 Oct 25]. ISIC | International Skin Imaging Collaboration. Available from: <https://www.isic-archive.com>
20. Mazhar T, Haq I, Ditta A, Mohsan SAH, Rehman F, Zafar I, et al. The Role of Machine Learning and Deep Learning Approaches for the Detection of Skin Cancer. *Healthcare*. 2023 Jan; 11(3):415. <https://doi.org/10.3390/healthcare11030415> PMID: 36766989
21. Hasan N, Nadaf A, Imran M, Jiba U, Sheikh A, Almalki WH, et al. Skin cancer: understanding the journey of transformation from conventional to advanced treatment approaches. *Mol Cancer*. 2023 Oct 6; 22(1):168. <https://doi.org/10.1186/s12943-023-01854-3> PMID: 37803407
22. Tadesse GA, Cintas C, Varshney KR, Staar P, Agunwa C, Speakman S, et al. Skin Tone Analysis for Representation in Educational Materials (STAR-ED) using machine learning. *Npj Digit Med*. 2023 Aug 18; 6(1):1–10.
23. Huang M, Yu W, Zhu D. An Improved Image Segmentation Algorithm Based on the Otsu Method. In: 2012 13th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing. 2012. p. 135–9.
24. Minaee S, Boykov Y, Porikli F, Plaza A, Kehtarnavaz N, Terzopoulos D. Image Segmentation Using Deep Learning: A Survey. *IEEE Trans Pattern Anal Mach Intell*. 2022 Jul; 44(7):3523–42. <https://doi.org/10.1109/TPAMI.2021.3059968> PMID: 33596172
25. Sezgin M, Sankur B. Survey over image thresholding techniques and quantitative performance evaluation. *J Electron Imaging*. 2004 Jan; 13(1):146–65.
26. Mutneja V. Methods of Image Edge Detection: A Review. *J Electr Electron Syst*. 2015 Jan 1;04.
27. Premana A, Wijaya AP, Soeleman MA. Image segmentation using Gabor filter and K-means clustering method. In: 2017 International Seminar on Application for Technology of Information and Communication (iSemantic) [Internet]. 2017 [cited 2023 Oct 26]. p. 95–9. Available from: <https://ieeexplore.ieee.org/abstract/document/8251850>
28. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely Connected Convolutional Networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [Internet]. Honolulu, HI: IEEE; 2017 [cited 2023 Aug 31]. p. 2261–9. Available from: <https://ieeexplore.ieee.org/document/8099726/>
29. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition [Internet]. arXiv; 2015 [cited 2023 Aug 31]. Available from: <http://arxiv.org/abs/1409.1556>
30. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016. p. 770–8.
31. Tan M, Le QV. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks [Internet]. arXiv; 2020 [cited 2023 Aug 31]. Available from: <http://arxiv.org/abs/1905.11946>
32. Shelke SK, Sinha SK, Patel GS. Development of complete image processing system including image filtering, image compression & image security. *Mater Today Proc*. 2023 Jan 1; 80:2167–71.
33. Ma L, Li M, Gao Y, Chen T, Ma X, Qu L. A Novel Wrapper Approach for Feature Selection in Object-Based Image Classification Using Polygon-Based Cross-Validation. *IEEE Geosci Remote Sens Lett*. 2017 Mar; 14(3):409–13.
34. Wang S, Tang J, Liu H. Embedded Unsupervised Feature Selection. *Proc AAAI Conf Artif Intell* [Internet]. 2015 Feb 10 [cited 2023 Oct 26];29(1). Available from: <https://ojs.aaai.org/index.php/AAAI/article/view/9211>
35. Regression Shrinkage and Selection Via the Lasso | Journal of the Royal Statistical Society Series B: Statistical Methodology | Oxford Academic [Internet]. [cited 2023 Aug 31]. Available from: <https://academic.oup.com/jrsssb/article/58/1/267/7027929?login=false>
36. Debelee TG. Skin Lesion Classification and Detection Using Machine Learning Techniques: A Systematic Review. *Diagnostics*. 2023 Jan; 13(19):3147. <https://doi.org/10.3390/diagnostics13193147> PMID: 37835889
37. LaValley MP. Logistic Regression. *Circulation*. 2008 May 6; 117(18):2395–9. <https://doi.org/10.1161/CIRCULATIONAHA.106.682658> PMID: 18458181
38. Noble WS. What is a support vector machine? *Nat Biotechnol*. 2006 Dec; 24(12):1565–7. <https://doi.org/10.1038/nbt1206-1565> PMID: 17160063
39. Breiman L. Random Forests. *Mach Learn*. 2001 Oct 1; 45(1):5–32.
40. Breiman L. Bagging predictors. *Mach Learn*. 1996 Aug 1; 24(2):123–40.
41. Schapire R.E. The strength of weak learnability. *Mach Learn* 5, 197–227 (1990). <https://doi.org/10.1007/BF00116037>

42. Wolpert DH. Stacked generalization. *Neural Netw.* 1992 Jan 1; 5(2):241–59.
43. Ho TK. The random subspace method for constructing decision forests. *IEEE Trans Pattern Anal Mach Intell.* 1998 Aug; 20(8):832–44.
44. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System [Internet]. 2016 [cited 2023 Aug 31]. Available from: <http://arxiv.org/abs/1603.02754>
45. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature.* 1986 Oct; 323(6088):533–6.
46. Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. In: Proceedings of the fifth annual workshop on Computational learning theory [Internet]. New York, NY, USA: Association for Computing Machinery; 1992 [cited 2023 Sep 1]. p. 144–52. (COLT '92). Available from: <https://dl.acm.org/doi/10.1145/130385.130401>
47. Hicks SA, Strümke I, Thambawita V, Hammou M, Riegler MA, Halvorsen P, et al. On evaluation metrics for medical applications of artificial intelligence. *Sci Rep.* 2022 Apr 8; 12:5979. <https://doi.org/10.1038/s41598-022-09954-8> PMID: 35395867
48. Mijwil MM. Skin cancer disease images classification using deep learning solutions. *Multimed Tools Appl.* 2021 Jul; 80(17):26255–71.
49. Jasil SPG, Ulagamuthalvi V. Deep learning architecture using transfer learning for classification of skin lesions. *J Ambient Intell Humaniz Comput* [Internet]. 2021 Mar 17 [cited 2023 Oct 25]; Available from: <http://link.springer.com/10.1007/s12652-021-03062-7>
50. Dubal P, Bhatt S, Joglekar C, Patil S. Skin cancer detection and classification. In: 2017 6th International Conference on Electrical Engineering and Informatics (ICEEI) [Internet]. Langkawi: IEEE; 2017 [cited 2023 Aug 24]. p. 1–6. Available from: <https://ieeexplore.ieee.org/document/8312419/>
51. Brinker TJ, Hekler A, Enk AH, Von Kalle C. Enhanced classifier training to improve precision of a convolutional neural network to identify images of skin lesions. Huynh, editor. *PLOS ONE.* 2019 Jun 24; 14(6):e0218713. <https://doi.org/10.1371/journal.pone.0218713> PMID: 31233565
52. Majtner T, Yildirim-Yayilgan S, Hardeberg JY. Optimised deep learning features for improved melanoma detection. *Multimed Tools Appl.* 2019 May; 78(9):11883–903.
53. Gupta S, R.J, Verma AK, Saxena AK, Moharana AK, Goswami S. Ensemble optimization algorithm for the prediction of melanoma skin cancer. *Meas Sens.* 2023 Aug; 100887.
54. Murugan A, Nair SAH, Kumar KPS. Detection of Skin Cancer Using SVM, Random Forest and kNN Classifiers. *J Med Syst.* 2019 Aug; 43(8):269. <https://doi.org/10.1007/s10916-019-1400-8> PMID: 31273532
55. Mustafa S, Dauda AB, Dauda M. Image processing and SVM classification for melanoma detection. In: 2017 International Conference on Computing Networking and Informatics (ICNI) [Internet]. Lagos: IEEE; 2017 [cited 2023 Oct 25]. p. 1–5. Available from: <http://ieeexplore.ieee.org/document/8123777/>
56. Linsangan NB, Adtoon JJ. Skin Cancer Detection and Classification for Moles Using K-Nearest Neighbor Algorithm. In: Proceedings of the 2018 5th International Conference on Bioinformatics Research and Applications [Internet]. Hong Kong Hong Kong: ACM; 2018 [cited 2023 Oct 25]. p. 47–51. Available from: <https://dl.acm.org/doi/10.1145/3309129.3309141>
57. Shalu, Kamboj A. A Color-Based Approach for Melanoma Skin Cancer Detection. In: 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC) [Internet]. 2018 [cited 2023 Oct 25]. p. 508–13. Available from: <https://ieeexplore.ieee.org/document/8703309>