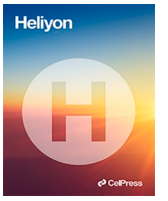
Heliyon 10 (2024) e31488

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Performance evaluation of E-VGG19 model: Enhancing real-time skin cancer detection and classification☆

Irfan Ali Kandhro [a], Selvakumar Manickam [b], Kanwal Fatima [a], Mueen Uddin [c], Urooj Malik [a], Anum Naz [a], Abdulhalim Dandoush [c,*]

[a] Department of Computer Science, Sindh Madressatul Islam University, Karachi, 74000, Pakistan
[b] National Advanced IPv6 Centre (NAv6), Universiti Sains Malaysia, Gelugor, Penang, 11800, Malaysia
[c] College of Computing and Information Technology, University of Doha For Science & Technology, 24449, Doha, Qatar

ARTICLE INFO

Keywords:
Skin cancer detection
Health care
Image segmentation
Pre-trained models
Machine learning and deep learning

ABSTRACT

Skin cancer is a pervasive and potentially life-threatening disease. Early detection plays a crucial role in improving patient outcomes. Machine learning (ML) techniques, particularly when combined with pre-trained deep learning models, have shown promise in enhancing the accuracy of skin cancer detection. In this paper, we enhanced the VGG19 pre-trained model with max pooling and dense layer for the prediction of skin cancer. Moreover, we also explored the pre-trained models such as Visual Geometry Group 19 (VGG19), Residual Network 152 version 2 (ResNet152v2), Inception-Residual Network version 2 (InceptionResNetV2), Dense Convolutional Network 201 (DenseNet201), Residual Network 50 (ResNet50), Inception version 3 (InceptionV3), For training, skin lesions dataset is used with malignant and benign cases. The models extract features and divide skin lesions into two categories: malignant and benign. The features are then fed into machine learning methods, including Linear Support Vector Machine (SVM), k-Nearest Neighbors (KNN), Decision Tree (DT), Logistic Regression (LR) and Support Vector Machine (SVM), our results demonstrate that combining E-VGG19 model with traditional classifiers significantly improves the overall classification accuracy for skin cancer detection and classification. Moreover, we have also compared the performance of baseline classifiers and pre-trained models with metrics (recall, F1 score, precision, sensitivity, and accuracy). The experiment results provide valuable insights into the effectiveness of various models and classifiers for accurate and efficient skin cancer detection. This research contributes to the ongoing efforts to create automated technologies for detecting skin cancer that can help healthcare professionals and individuals identify potential skin cancer cases at an early stage, ultimately leading to more timely and effective treatments.

1. Introduction

In the United States, the Skin Cancer Foundation estimated in 2022 that more than 9500 people receive a skin cancer diagnosis per day, also more than two people die from the condition every hour [1–3], the skin can be affected by a wide variety of tumor types,

☆ Open Access funding provided by the Qatar National Library.
* Corresponding author.
E-mail address: Abdulhalim.dandoush@udst.edu.qa (A. Dandoush).

https://doi.org/10.1016/j.heliyon.2024.e31488
Received 4 May 2024; Accepted 16 May 2024
Available online 17 May 2024

Abbreviations

BCC	Basal Cell Carcinoma
SCC	Squamous cell carcinoma
CPS	Computerized Pathology Systems
ABCD	Airborne Contact Dermatitis
CAD	Computer-Aided Diagnosis
CNN	Convolutional Neural Network
KNN	k-Nearest Neighbors
SVM	Support Vector Machine
E-CNN	Enhanced Convolutional Neural Network
ANN	Artificial Neural Network
SVM	Support Vector Machine
ML	Machine Learning
DL	Deep Learning
ISIC	International Skin Imaging Collaboration
HOG	Histogram of Oriented Gradients
PSO	particle swarm optimization
DCNN	Deep Convolutional Neural Networks
PCA	Principal Component Analysis
GLCM	Gray Level Co-occurrence Matrix
UV	Ultraviolet

making skin cancer the most prevalent type of cancer [2]. The most prevalent types of skin cancer, in addition to melanoma, are basal cell carcinoma (BCC), squamous cell carcinoma (SCC) [3–6], and epithelioma. A thorough recovery and the avoidance of the cancer metastasizing into deeper tissues depend on prompt identification. However, there are currently few non-invasive, economically viable approaches for broadly aiding diagnosis and characterizing skin lesions objectively. Typically, when suspicious cases arise, dermatologists conduct visual examinations of the skin using epiluminescence [3–5,7,8]. Skin cancer refers to a class of diseases where aberrant skin cells proliferate out of control, leading to the development of tumors. Most of these tumors are caused by UV (ultraviolet) radiation exposure and unprotected skin damage [2–4]. Basal cell and squamous cell carcinomas make up most skin cancer cases (99 %), whereas melanomas account for barely 1 % [2]. Since the 1970s, there has been a rise in the incidence of skin cancer, and there are many different methods used by doctors to identify this disease [3–5,7]. Trained clinicians frequently use specific criteria, such as visual inspection, dermoscopy, and biopsy, to diagnose the malignancy of a lesion. Image identification accuracy has increased significantly due to dermoscopy, going from 75 % to 84 %. For those with skin disorders, it might be difficult and time-consuming to make an accurate diagnosis, and a doctor's knowledge is still necessary [9]. In circumstances when trained professionals are not easily available, computer-assisted diagnosis can analyze dermoscopy operations, eliminating variances between samples and practitioners. However, previous computer-assisted skin image categorization systems faced two main challenges: inadequate data and inconsistent image quality [9–12]. The use of substantial preprocessing, segmentation, and feature extraction has been replaced using machine learning (ML) technology in modern techniques. This strategy saves time and effort and is consistent with the wider trend of using ML approaches for cancer detection [13]. Machine learning algorithms have increased the accuracy of cancer by 16 %–25 % throughout the previous 20 years. Among these techniques, convolutional neural networks (CNNs), a type of deep learning, stand out as a reliable and popular approach for image recognition and classification that makes use of advanced computational algorithms and big datasets. In fact, researchers have reduced the frequent use of traditional ML techniques that demand extensive background knowledge and lengthy preparation stages. Skin cancer in photos may be accurately identified by deep learning-based classifiers, outperforming dermatologists in this regard. They achieve this by using CNNs to support the creation of dermatologists' computer-aided classification systems for skin lesions. However, a lack of sufficiently annotated and described photos of rare classes hinders the availability of high-quality medical image training sets. Smaller data sets have a lower success rate for standard CNNs. Furthermore, some researchers use unusually deep CNN models, such as Resnet152 [14] with its 152 layers, that raises concerns regarding processing costs for therapeutic applications. These models increase network classification performance but need a significant amount of processing power. To minimize data overfitting, researchers used CNNs trained to categorize skin lesions and supplement them with characteristics extracted from real-world picture datasets such as ImageNet [11]. This integration of AI, particularly Computer-Aided Diagnosis (CAD), dramatically simplifies and reduces the cost of identifying and treating tumor disorders. Skin lesions were originally recognized visually, even though detecting organ issues frequently involves imaging techniques like MRI, PET, and X-rays. This has now been made possible by several techniques, including CT, dermoscopy image processing, and clinical screens. Dermatologists with less training are more likely to misdiagnose skin lesions, mainly because it takes time and is frequently subjective and imprecise to evaluate and analyze photographs of lesions. Accurately capturing skin disease on camera is difficult. The capacity of CAD and Computerized Pathology Systems (CPS) to detect skin cancer has been greatly improved because of the incorporation of machine learning. Lesion image classification and image preprocessing are important steps in this procedure. However, there is an urgent requirement to increase skin cancer survival rates and outcomes by early identification. To speed up and improve AI in medicine is

becoming more and more popular among doctors as a tool for diagnostic decision-making [13,15–20]. In this study, we provide a fresh approach to identifying benign from malignant melanoma, a vital issue in skin cancer classification.

Benign tumors consist of normal cells that proliferate and divide without impairing the normal functions of surrounding cells. Melanoma tumors, on the other hand, are composed of aberrant cells that exhibit excessive growth, disrupting normal cell functions. These cells undergo uncontrollable proliferation and division [4,5,21]. Fig. 1 illustrates the comprehensive classification of skin cancer.

Our research provides the following significant contributions.

- Present an enhanced version of the VGG-19 model for the detection of skin cancers, incorporating additional layers and hyperparameters.
- Employ robust feature selection and extraction methods to derive deep features from skin cancer images.
- Compare the performance of the enhanced VGG-19 model (E-VGG19) with other architectures that combine machine learning and pre-trained models.

The remainder of the paper is organized as follows: In Section 2, we give a review of the literature on skin cancer categorization, encompassing different ML and DL models. The approaches used are examined, with an emphasis on both human and automated feature extraction methods. The architecture we propose is described in Section 3 along with further details on the approaches of our DL and ML models used. The results of the experiment from the distinct DL, ensemble models, and ML are presented in Section 4 of this paper. Additionally, a thorough comparison between our suggested model and skin cancer detection models is provided in this part. Lastly, in Section 5, we conclude the paper and present potential directions for future research in this domain.

2. Related work

In recent years, skin cancer detection has become a critical area of research due to its significant impact on public health. Deep learning techniques, especially pre-trained CNN models, hold promise for enhancing accuracy and efficiency in skin cancer detection, segmentation, and classification. In this section we will discuss the related works to the four main techniques used in this field. We start by the literature working on features extraction in subsection A. Then we move to discuss the works related to the features selection and fusion in subsection B. After that we explore the machine learning classifiers used to detect the skin cancer. Last, the CNN-based DL techniques are discussed in subsection D.

2.1. Features extraction

The feature extraction process of the present research included the use of ABCD to extract general features. With respect to color features, they were obtained from HSV, LAB, and RGB. Textures were obtained through Gray-Level Co-occurrence Matrix (GLCM). For selecting important factors and using them in a mathematical model later on, a genetic algorithm was used [22]. A combination of first-order moments and color histogram features was proposed by the researchers who favored a superior approach in feature extraction; this would entail integration with Histogram of Oriented Gradients (HOG) technique. This fusion of image and text was first introduced as a way to enhance the traditional bag-of-words approach [23]. In their paper, the researchers noted that the utilization of these different feature extraction methods together might result in better quality features that represent more accurate key points.

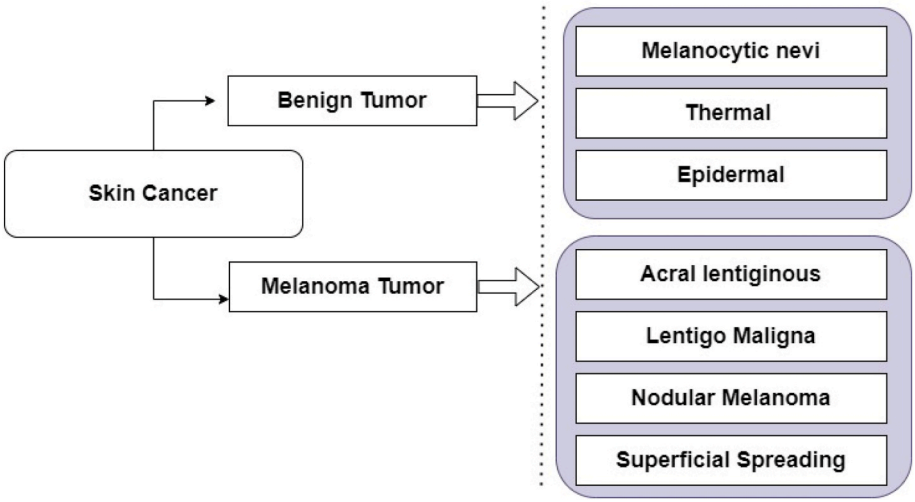


Fig. 1. The detailed classification of skin cancer.

Researchers have discovered eight links to nearby objects with the use of LTRP, a special tool. In addition to this, it employed Local Binary Patte-rn and also Local Vector Pattern. These were able to locate the conne-ctions. The researc-hers als-o employ CST in describing these connections. Using LBP + LVP worked better as compared to other methods [24].

The researchers employed the ABCD approach as a means of gathering an array of symmetry-related characteristics from the lesions. This comprehensive method delves into parameters such as border features, color attributes, and diameter measurements, providing a nuanced understanding of the lesions' structural intricacies [25]. In tandem with this, the classification SVM leverages the GLCM and HOG techniques. These methods, operating synergistically, play a crucial role in extracting the textural characteristics inherent in the lesions.

The analysis of second-order texture characteristics is not limited to pixel values only. The texture features under examination are carefully chosen to embody the most significant aspect in the picture. A strategic approach, as such, allows deeper exploration into the subtle textural attributes of these lesions thereby contributing towards a comprehensive characterization and classification process. ABCD, GLCM, HOG and SVM classified systems' integration also signifies that this research embraces multiple pronged approaches aimed at increasing accuracy and depth of lesion analysis [25]. The colour histogram uses a GLCM spatial and colour information for OPP, HSV and RGB colour spaces [26]. To determine local features of a sample Speeded Up Robust Feature calculates global textures of the sample amongst other factors that contribute in extraction.

In the classification phase, SVM performed better than K-Nearest Neighbors (KNN), whereas SURF features beat GLCM and SIFT [27]. The Local Binary Pattern extracts the statistical features of mean texture, smoothness, skewness, and energy [99]. They blended color and texture elements using a bag of words method during the feature extraction stage. Additionally, to highlight the colour information, HL and HG were bagged independently and mixed with other bagged Zernike and bagged angles of colour vector [28]. The standard deviation, minimum, mean, skewness, and kurtosis for each component (R, G, and B) were calculated to extract color-based features, whereas the wavelet transform was used to produce texture-based features [29]. ReliefF, Chi2, RFE, and CFS were the four distinct feature selection methods that the researchers investigated. After that, they used standard score transformation to feature normalization [30].

2.2. Features selection and fusion

The researchers offered numerous ways for extracting the most useful information using machine vision to classify skin cancer. They integrated first-order histogram features and GLCM used PCA to minimize dimensionality [31]. Furthermore, the researchers used the cumulative level difference mean to modify ABCD and pick important features using the ranking and selection technique of features via the Eigenvector [32]. To identify a melanoma, they performed network training on the important characteristics, notably those obtained by the GLCM technique, and optimized it utilizing a binary bat technique, which gives a meaningful collection of features [33]. This unique strategy, referred to as the optimized framework by fusing PCA with features with higher entropy values [34]. The major objective is to use an accurate feature extraction approach to discriminate between malignant and non-cancerous tumors. The researchers propose combining a slew of features with accelerated, reliable features [35]. However, more optimum features must be extracted and optimized further to lower the error rate associated with complicated features.

2.3. Machine learning classifiers

In the area of machine learning-based skin cancer detection, researchers have achieved impressive results. An SVM was used to analyze melanoma skin cancer using one version of the ISIC dataset, yielding a 96.9 % accuracy rate [36]. They also employed Several classifiers to diagnose melanoma, including Decision Tree. KNN, SVM, Ensemble. The resulting accuracy was 89.5 %, 86.5 %, 76.0 %, and 100 % [37]. Another technique entailed developing a skin lesion categorization system based on decision trees that obtained an accuracy of 97 % on both the HAM 10,000 and ISIC 2017 datasets [38]. The Nave Bayes classifier was used in conjunction with the Dermatology Information System and DermQuest to attain a precision of 97.6 % [39]. They also investigated other classifiers and their versions, with the SVM form attaining the greatest accuracy of 83 % [40]. The Nave Bayes classifier detected melanoma, keratosis, and benign skin diseases with accuracies of 91.2 %, 92.9 %, and 94.3 %, respectively [41]. Another study suggested a method for melanoma detection that used a KNN classifier based on fuzzy decision ontology, reaching an accuracy of 92 % on the DermQuest and Dermatology Information method datasets [42]. Finally, employing ANN and SVM classifiers, researchers created a system for melanoma identification with accuracy of 96.2 % and 97 %, respectively [43].

Recent advancements in computer vision and deep learning have considerably advanced the science of prediction. These advancements have had a significant influence on a wide range of computer vision applications, including cancer detection, autonomous driving, medical picture segmentation, and others [44–46]. However, since large, well-annotated datasets are scarce for this endeavor, skin cancer detection confronts a unique obstacle. Current techniques to skin cancer classification may be divided into two categories: those that use handmade features and those that use deep learning for automated feature extraction.

Mahmoud and Elgamal [47] used methods for feature extraction, notably wavelet processing. They retrieved these traits, and then used dimensionality reduction techniques before feeding them into machine learning models. They used k-nearest neighbor (k-NN) and artificial neural networks (ANN) to classify skin cancer based on clinical observations, detect features within Dermoscopic pictures, and quantify tumor depth. Approximately 81 descriptors, including color, texture, form, and pigment network properties, were retrieved. They reached an amazing 95 % overall accuracy by combining logistic regression and neural networks.

In another study [48], grayscale morphology was used to reduce information distortion caused by undesirable objects in the photographs, such as hair particles. To segment lesions, three unique methods were developed: a 3-D colour clustering idea, dynamic

thresholding, and global thresholding. Moreover, original photos were converted from the RGB scale to the HSI space to compute basic colour characteristics. The most accurate supervised learning system employed to classify skin cancer has a 77.6 % overall accuracy. The categorization of melanoma skin cancer was suggested in Ref. [48], with the classifier being support vector machines (SVM). Prior to feature extraction, images were preprocessed and segmented using thresholding. The Gray Level Co-occurrence Matrix (GLCM) was used to extract features, which focused on four Dermoscopic characteristics symbolized by the acronym ABCD, which stands for " Diameter, Asymmetry, Colour, Border."

Furthermore, the Principal Component Analysis (PCA) method was used to reduce the dimensionality yielding a remarkable total accuracy of 92.1 %. In later study [49] multiple supervised learning algorithms for the categorization of benign and malignant melanoma pictures were examined and analyzed. We compared the algorithms SVM, ANN, and AdaBoost. In order to analyze it further, some pre-processing methods were employed such as de-noising, image sharpening, resizing and segmentation. Segmentation was done using K-means clustering technique with best results being obtained from AdaBoost and SVM algorithms on the experimental dataset. Use of deep learning (DL) in automating feature extraction in skin lesion classification has been gaining popularity among researchers. For instance Ref. [50], used Ph2 dataset which consists of about 200 skin cancer images for training purposes only. To solve the problem of a small dataset, they did so by performing data augmentation techniques where they manipulated their photos thereby increasing their number to 6600. This involved updating the neural network's weights based on AlexNet architecture that had previous knowledge before applying stochastic gradient descent (SGD) algorithm to update the net weights again. They tested it using metrics like accuracy, precision; recall; and specificity thus exhibiting that this model outperformed the previously existing ones.

In contrast, Zhang et al. [50] focused on limitations relating to DCNNs in skin lesion categorizing based on deep convolutional neural networks including lack of attention towards semantically important regions within an image. To overcome this issue, they developed an attention residual learning CNN composed of several attention residual learning blocks. They put this network through its paces on the ISIC-skin 2017 dataset and discovered that it outperformed state-of-the-art approaches. Another strategy required the creation of a melanoma classification ensemble of neural networks [51]. Thermoscopic datasets of several races were used, and the categorization considered numerous parameters such as colour, border, and texture, the Deep convolutional neural networks were used by the researchers in Ref. [50] for autonomous melanoma identification. They employed residual learning to avoid overfitting and used deep convolutional network and the residual network to create a two-stage network. The proposed framework achieved an area under the curve (AUC) score of 80.5 % when tested on the ISBI 2016 Skin Lesion Analysis Challenge. Another study [51] built many using the HAM10000 dataset to train neural network designs beforehand, which contains many forms of lesions of the skin. The ResNet50 model produced the greatest accuracy, they discovered. They used an ensemble strategy, integrating ResNet50, VGG16, and DenseNet, to increase accuracy even more. They were therefore able to attain an accuracy of 84 %.

Using seven classes from the HAM10000 dataset, Chaturvedi et al. [52] built on this work by employing deep convolutional neural networks for multi-class skin cancer diagnosis, obtaining good accuracy for both individual and ensemble models. Traditional machine learning algorithms like logistic regression and Support Vector Machines (SVM) can still produce excellent results when given the right features and image pre-processing methods, even though deep learning is known for its automatic feature extraction abilities. These conventional models are a good substitute for deep models since they frequently use fewer processing resources. The experimental findings from the proposed study provide support for the above assertion.

2.4. CNN based techniques

In the area of machine learning-based skin cancer detection, researchers have achieved impressive results. The CNN with SVM classifer was used to analyze melanoma skin cancer using the ISIC dataset, yielding a 97.8 % accuracy rate [36]. Studies on the detection of skin cancer using medical image analysis have advanced significantly in recent years [53] The skin cancer diagnosis standard was enhanced during the ISIC 2018 event through a challenge competition. In each of these attempts, the researchers have used various classification techniques and algorithms to try to improve the diagnosis accuracy. Its architecture underwent several significant changes after CNNs were added to the image classification tasks. CNNs are thought to be the most cutting-edge method for classifying images because they mimic human visual cognition. We concentrate on deep learning methods for analyzing skin cancer pictures in our literature review, even if there is much research on image categorization in the literature.

Over the past few decades, numerous research has been undertaken that suggests extremely successful image classification algorithms for the aim of skin cancer diagnosis. The binary classifications of malignant keratinocyte benign seborrheic keratoses vs. carcinomas and benign nevi vs. malignant melanomas were used to train a CNN model [27,52–54] using a collection of clinical pictures. The most prevalent malignancies are searched after in the first instance, while the deadliest skin cancers are sought after in the second. Their suggested model performs equally well in both tasks as did the entire group of experts tested. This demonstrates that smart computer-based algorithms can classify skin malignancies with accuracy on par with dermatologists. A selection of techniques for the data-limited verification of training and learning were described in Refs. [55,36,50–52]. When the network gets deeper, overfitting and degradation become problems, so they use residual learning to address them. They also developed a fully convolutional residual network for the precise segmentation of skin lesions, and they enhanced its capabilities by incorporating a multi-scale contextual information integration scheme. Instead of using the entire dermoscopy images, their architecture eliminates the shortage of training data by enabling a more accurate and representative extraction from the segmented outputs. Dermoscopic images and the diagnoses that go along with them are used.

The Google's Inception v4 [15,19] CNN's architecture was trained and verified. Additionally, for clinical data on two levels, they used a set of 100 images. The first level took into consideration dermoscopy pictures. Their method's main output is a binary diagnostic classifier that uses dermoscopy melanocytic images as its basis. The work in Ref. [13] classified skin cancer using a related ECOC

SVM-based classifier and an AlexNet CNN-based classifier. As compared to, their model exhibits greater specificity, accuracy, and sensitivity when considering a four distinct forms of skin tumors are depicted in a series of 3753 images [13,56,16,21,50,52–54].

For the greatest early skin cancer diagnosis, the works [17,57] used CNN. They optimized the CNN parameters using IWOA. The network's weights and biases were best chosen using the optimization algorithm to minimize the intended output mistakes. In this paper [18,22,23,49] author proposed a technique for feature optimization using a variant of the PSO based on dermoscopic images.

Table 1
Summarized Skin Cancer Detection with Deep Learning and Machine learning Techniques.

Ref	Skin Cancer Detection	Dataset	Training and Classifier Algorithm	Description	Results
[62]	Benign/malignant	ISIC dataset (2016)	A very strong remnant of CNN and FCRN	For precise lesions segmentation, FCRN combined with a multi-scale contextual information integration approach was developed	Jaccard index (82.9), dice coefficient (89.7), specificity (95.7), sensitivity (91.1), and accuracy (94.9)
[63]	Melanoma/SK	ISIC dataset (2017)	Multi-scale depth CNN	The Inception-v3 model, which was trained on ImageNet, was employed in the suggested model	AUC (94.3), Accuracy (90.3)
[64]	Malignant melanoma/SK	ISIC dataset (2017)	AlexNet, ResNet-18, and VGG16 are three previously trained deep models that were used to extract features for SVM classification.	Logistic regression function was used to convert SVM scores to probabilities for evaluation.	Average AUC (90.69)
[65]	AK/benign keratosis/nevus/ vascular lesion/and dermatofibroma	ISIC dataset (2016, 2017)	CNN model using LeNet methodology	To improve system performance, the adaptive piecewise linear activation function was employed.	(95.86) Accuracy
[66]	Atypical nevi/ dermatofibroma, benign keratosis/melanocytic nevi/ BCC/AK/IC/and vascular lesions	PH2 and HAM10000 dataset	Deep CNN architecture (ResNet 152, DenseNet 201, Inception v3, and v2) InceptionResNet	In terms of overall mean results, deep-learning models outscored highly skilled dermatologists by at least 11 %.	ResNet 152: 98.61-98.04; DenseNet 201: 98.79-98.16; Inception v3: 98.60-97.80; InceptionResNet v2: 98.20-96.10)
[67]	Benign keratosis, Bowen's disease, BCC, melanocytic nevus, Bowen's disease, AK, vascular lesion, and dermatofibroma	ISIC dataset (2018)	PNASNet-5-Large, SENet154, InceptionV4, and InceptionResNetV2	To set up network parameters and fine-tune them, an image-net model that had been trained was utilized.	Score for Validation (76)
[68]	BCC/non-BCC	FF-OCT images: 297	ResNet18 pruned	It was determined, using K-fold cross-validation, the effectiveness of the suggested system	(80) Accuracy
[69]	Melanocytic nevi, BCC, SK, and SCC	1300 pictures of skin lesions	CNN	Mean subtraction and Pooled Multiscale with averaging in augmented feature space.	Accuracy (81 %)
[70]	Lupoma, fibroma, scleroderma, and melanoma	ISIC dataset (2016)	Clustering using deep region-based CNN and fuzzy Logic.	A combination of fuzzy C-means and region-based CNN achieved greater accuracy in disease identification	Precision (94.8) responsiveness (97.81) precision (94.17) (95.89) F1_score
[71]	Benign/malignant	1760 dermoscopy pictures	5-fold cross-validation on CNN	Images were prepared based on melanoma cytological results	Specificity (88.1), sensitivity (80.9), and accuracy (84.7)
[72]	Combination nevus and malignant melanoma	Dataset (Derma, Dermnet, Danterm and DermQuest)	CNN	For fine-tuning, the pre-trained BVLC-AlexNet model from the ImageNet dataset was employed.	Average mean precision (70)
[73]	Malignant/benign	ISIC 2016 and MED-NODE datasets	Six layers thick CNN	Image illumination had an impact on how well the system performed.	Precision (77.50)
[74]	Nonmelanoma/melanoma	ISIC dataset (2016), ISBI (2016)	Fully hybridized CNN with autoencoder, decoder, and RNN	The proposed model performed better than the most advanced SegNet, FCN, and ExB architecture	Precision (98) Sensitivity (95), specificity (94), and Jaccard index (93)
[75]	Benign/malignant	ISIC dataset (2017)	CNN with two layers and a new regularized	The suggested regularization strategy reduced complexity by imposing a penalty on the classifier's weight matrix's dispersion value	Accuracy 97 %, Specificity, and Sensitivity 94 %.
[76]	Cancerous vs. benign	ISIC dataset (2016,2017)	Dark CNN	Data balance required data augmentation.	Precision (81), accuracy (80.3), and AUC (69)

Along with attraction operations, flee and sub-swarms, local and global food and enemy signals, and mutation-based local exploitation, their proposed algorithm also includes a variety of matrix representations to prevent the primary PSO algorithm from convergent too early. To diversify the search processes, they also used probability distribution and dynamic matrix representations. They use SVM and KNN as two classifiers to categorize skin lesions. Authors in Ref. [19] suggested Skin cancer classification using an automated hyper-parameter CNN [1,5,9,27,31–41,21,57–62]. The suggested method makes use of CNN's hyperparameters by choosing the right encoding schemes. The deep larning and machine leaning methods,descriptions of models, datasets, and results on lesions of skin cancer is summarized in Table 1.

3. Proposed E-VGG19 model and Methodology

The architecture of proposed Enhanced VGG19 models is thoroughly explained in this section with comparison of other pretrained models. The basic structure E-VGG19 model is based on CNN, The CNN is a fundamental deep learning model that can be used to assign weights and biases that can be trained to an image's input to determine its relevance to various objects and aspects of the image and distinguish one from the others. During the training phase of the main CNN model, to understand the properties of different training patterns, the filters are manually developed. The Visual Cortex organization served as inspiration for the CNN model's architecture, which is strikingly comparable to the connection pattern's architecture of human brain neurons. Individual neurons only react to stimuli in a small area of the visual field known as the receptive field. These fields should be combined to cover the entire visual region.

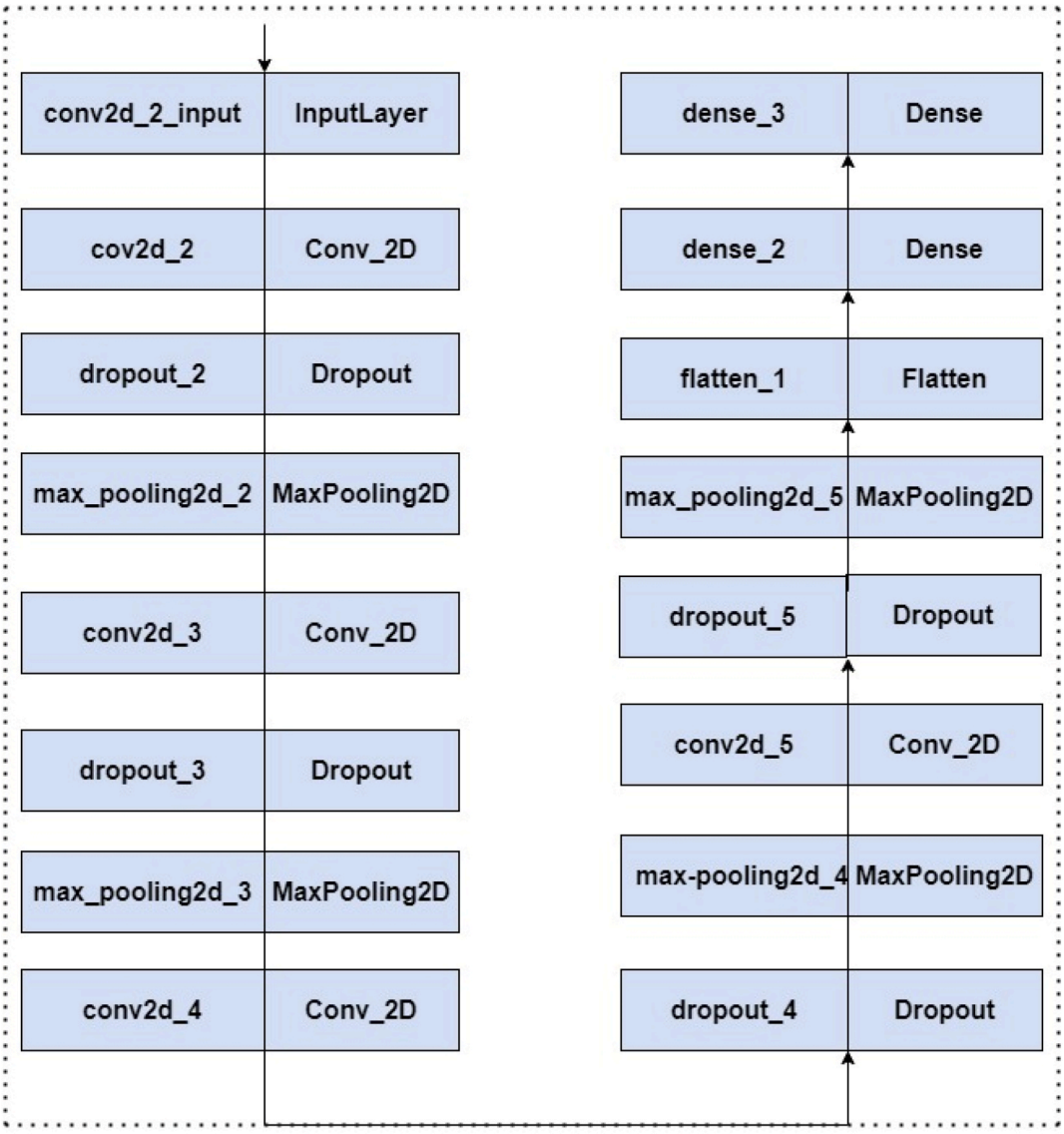


Fig. 2. Configuration of E-VGG19 architecture A. Image segmentation.

In addition, Machine learning techniques are integrated with the CNN based model (i.e., E-VGG19) in our proposal to enhance overall performance and robustness in skin cancer detection. The E-VGG19 model that we explain hereafter serves as a feature extractor to feed a separate ML classifier (e.g., Support Vector Machine and Random Forest) as shown in Fig. 3.

The E-VGG19 network has numerous parameters and layers, as mentioned in the previous subsection, which can make recognition tasks more difficult. VGG19 is a CNN model with an intricate 19-layer architecture. The pretrained VGG19 model is a fundamental component of our work, having been first trained on the 1000 class ImageNet dataset [42]. We investigate the VGG19 architecture and carry out experiments with fine-tuning in every layer. A set of four fully connected layers is used in place of the top layers to customize the model for our dataset. By strategically retraining the VGG19 architecture, we hope to better suit its capabilities for sophisticated feature extraction and classification by adjusting it to the nuances of our dataset. In our method, a convolutional layer was removed from the third block while in the first three blocks of VGG19, the settings were maintained constant. A BN layer was also added after each pooling layer. To replace the fully connected layer, a GAP layer was finally used. We ignored the dropout layer and substituted a BN layer even though many CNNs use dropout layers. The main justification for doing away with the variance offset those results is the dropout layer, which might help to improve the accuracy of the results' mean and standard deviation.

As a result, during the training phase, there might be a shift in the mini-batch sample distribution. arbitrarily, which could cause gradients to vanish or explode. This scenario has an impact on the layers below it as well. The input map was normalized and then the network was fitted using batch normalization. To produce a zero means and a unique variance to solve this problem After convolution, CNNs are made up of two enormous, fully connected layers with enormously more parameters. However, this circumstance may make the network more computationally taxed and slow its performance. As a result, overfitting may occur when the parameters are too many. Releasing itself as a vector that is only one dimension, the fully connected layer runs the risk of losing information, whereas the global average pooling layer gathers feature information by combining the most recent to gain from preserving feature information, input a feature map into a whole. Therefore, rather than using fully connected layers, the suggested method in this work makes use of a global average pooling layer. A global average pooling layer also contains fewer parameters than a fully connected layer, which could speed up calculations and satisfy the demands of real-time diagnosis [77]. Fig. 2. The network architecture of the recommended method is displayed.

The objective of the skin lesion segmentation stage is to accurately determine the border of a skin lesion. By accurately localizing the lesion boundaries, clinicians can gain insights into the extent and shape of the lesion, which is crucial for diagnosis and subsequent medical decisions. The precision of this stage holds paramount importance, as the lesion boundary serves as the foundation for numerous parameters utilized in assessing the risk of melanoma. In this context, we implement a lesion segmentation technique based on texture distinctiveness to precisely identify lesion boundaries. The skin lesion segmentation algorithm initiates by acquiring representative texture distributions and determining the texture distinctiveness measure for each distribution (refer to Fig. 4). Subsequently, a texture vector is extracted for every pixel in the image. Utilizing this set of texture vectors as a foundation, a Gaussian mixture model is then applied to effectively learn the underlying texture distributions.

Lastly, the texture distinctiveness metric is used to quantify how different a texture distribution is from all other texture distributions. Flow chart illustrating the combination of the texture distinctiveness map and the starting regions. In the illustration of the early regions, every solid hue designates a particular region. The pixels in the picture are categorized as belonging to the lesion class or to the normal skin in the second stage. This is accomplished by segmenting the image into multiple areas. To locate the skin lesion, these regions are integrated with the texture distinctiveness map. The geometric feature of the melanoma skin lesion is one of its primary characteristics. Therefore, we suggest extracting the segmented skin lesion's geometric features. The following common

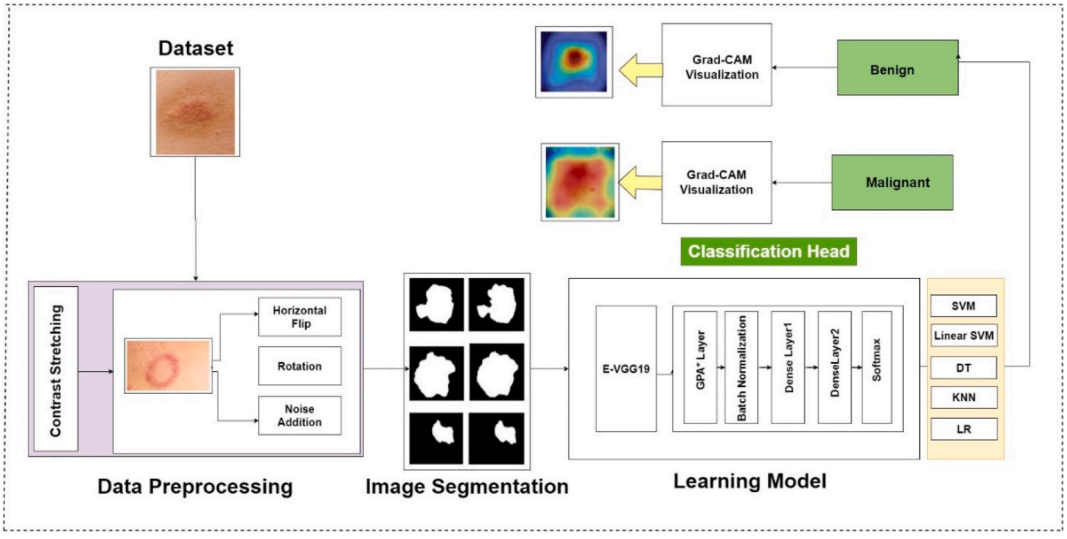


Fig. 3. Proposed (E-CNN) architecture of skin lesion detection and classification.

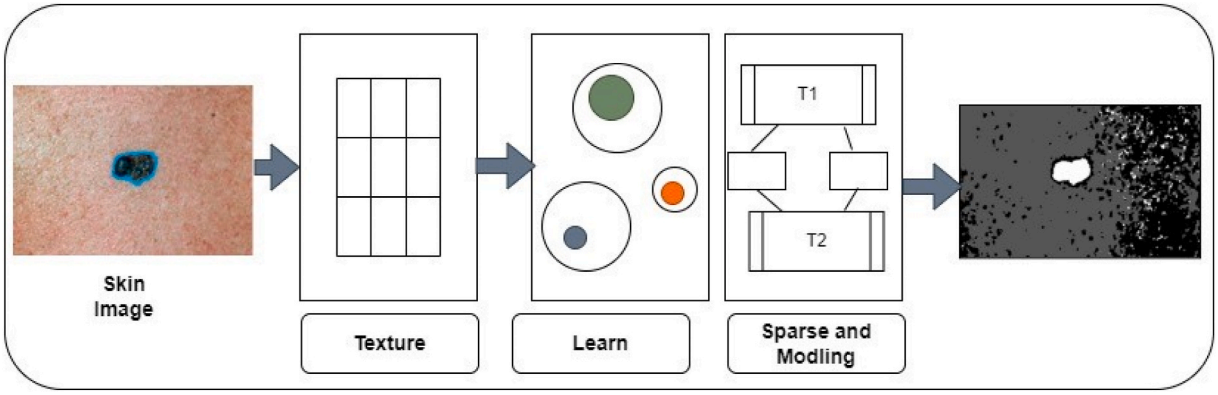


Fig. 4. Process showing texture distinctiveness map and initial regions.

geometry properties were employed in this instance: Area, Perimeter, Greatest Diameter, Circularity Index, and Irregularity Index, all taken from.11 To extract the geometrical features of the skin lesion, the image blob of the segmented picture that only contains the lesion is analyzed. The following are the many features that were extracted: Area (A): The lesion's pixel count. Number of edge pixels is the perimeter (P). Major Axis Length, also known as Greatest Diameter (GD), is the length of the line joining the two farthest boundary points and passing through the lesion centroid as mentioned Equations (1) and (2), where (xi, yi) is the position of the ith lesion pixel and n is the number of pixels inside the lesion. The length of the line that connects the two closest boundary points and passes through the centroid of the lesion blob is known as the minor axis length, or shortest diameter (SD). The Circularity Ratio (CRR): It provides consistency in the shape.

$$(xc, yc) = \left(\frac{\sum_{i=1}^{n} xi}{n}, \frac{\sum_{i=1}^{n} yi}{n}\right) \tag{1}$$

$$CRC = \frac{4A\pi}{p2} \tag{2}$$

3.1. Machine learning

In computer science, machine learning enables computers to learn from data without explicit programming. Numerous machine learning approaches, like SVM and logistic regression, are popular machine learning approaches frequently employed for classification and regression issues. The foundation of binary SVM is the creation of a maximum geometric and functional margin by using a separating hyperplane to divide the data points into two groups. But another significant linear machine learning classifier that uses a preset threshold to aid with the binary classification problem is logistic regression, which makes use of the sigmoid function. Using logistic regression or linear kernel support vector machines (SVM) usually involves certain trade-offs, depending on the properties and attributes of the reference dataset. In this case, we use both linear kernel support vector machines (SVM) and voting for the ultimate categorization of skin cancer will be done using logistic regression. To improve feature representation, we first use local binary pattern histogram and contourlet transform image processing methods before using these machine learning classifiers for the diagnosis of skin cancer. Following dimensionality reduction with PCA, the characteristics obtained by these two image-processing techniques are fed into the ML classifier. In this part, we provide a quick explanation of PCA and these image processing techniques. While the feature extraction procedure is fully automated, DL approaches prevent us from manually removing characteristics from the photos [31]. Deep learning thus suffers when handmade features have the potential to improve accuracy. ML models enable the provision of features following a thorough feature extraction process.

3.2. Transfer learning

With transfer learning [45], we can apply the best-performing models to the most widespread dataset, "ImageNet". Additionally, it allows us to modify the model according to the needs of the problem and the domain. Once the input size was adjusted, we used the pre-trained model/problem straight away, substituting the required number of output classes (two in our instance) for the original ImageNet problem's 1000 classes. In addition, we have included a few more layers to refine the model for the problem of skin cancer detection. An alternative to using a pre-trained model directly is the deep learning model's training from the beginning. As it can be seen in the results section, compared to the model created using ImageNet weights, the deep learning model was built from scratch and outperformed on the dataset. For this study, we tested five strong deep neural network architectures, selecting the top-performing model to build an ensemble using our suggested machine-learning technique. The CNN models are trained using a learning rate of 0.0001, Adam is used as an enhancer, and a few modifications is recommended for the models being considered for skin cancer diagnosis. The input photographs are scaled to 224 × 224 pixels during the preprocessing step to ensure uniformity in dimensions. The

skin cancer detection model, known as E-VGG19 and shown in Fig. 3, has an architecture that includes layers that are well-suited to different tasks. The E-VGG19 model intentionally uses the convolutional layers first to extract complex information from the input images. Convolutional layers are important because they let the model capture hierarchical representations, which in turn helps it identify subtle patterns associated with skin diseases. Subsequent pooling layers help retain important information while lowering

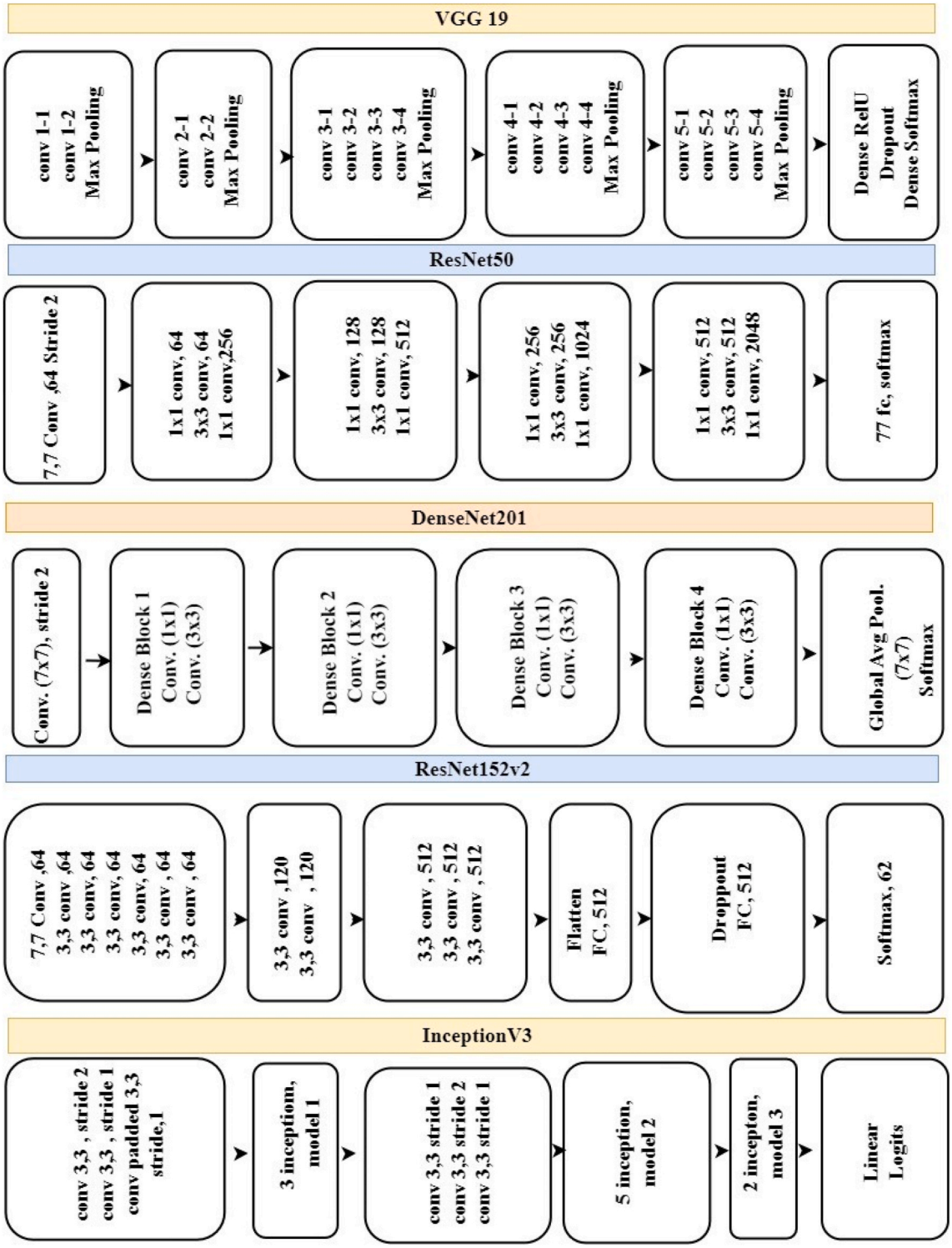


Fig. 5. Architecture of Pretrained models.

computational complexity by contributing to spatial reduction. The specific architecture can vary depending on the dataset, model complexity, and available computing resources. However, the goal is to achieve accurate and reliable skin cancer detection, which might help with early care and diagnosis [52,78].

3.3. VGG19

VGG19 is a pretrained CNN model with 19 layers, as the name would imply. Initial training of the trained VGG19 [42] architecture was conducted using 1000 classes from the ImageNet dataset. The VGG19 architecture is tested and retrained on our dataset in this work by fine-tuning each layer individually and substituting the top layers with four fully connected layers. The classification result is then obtained by applying the sigmoid layer. Resizing the input to match the VGG19's original input size makes it compliant.

3.4. ResNet50 [11,19,66]

Retraining the ResNet50 [19] architecture using the dataset makes it possible to test and improve across all levels. With "relu" serving as the activation function, four fully connected layers take the role of the upper layers. After that, a sigmoid layer is used to interpret the data into two diagnostic groups. In this case, identity mapping is used by the network to train far deeper networks and solve the vanishing gradient problem. Although, this identity mapping lacks parameters the output of the layer below is added to the layer above. The outputs from earlier levels are added to the outputs of stacked layers when layer connections are skipped [11]. With their pretrained weights, the ResNet50 architecture is maintained like the earlier pre-trained CNNs, At the top, four completely connected layers have been added, with "relu" serving as the activation function. The last output layer's activation function is the sigmoid function. We also implemented ResNet152v2, which has 152 layers in the model and a comparable architecture to ResNet50 and subjected it to thorough testing.

3.5. InceptionResnetV2 [11,24,66]

The ResNet models and Inception were used to create this model [20,24,66]. On our dataset, the InceptionResNetV2 architecture is retrained, and adjustments are recommended for every layer. To categorize the picture into two diagnostic groups, a sigmoid layer is applied at the end, and four completely linked layers are put in place of the top layers. The ISIC dataset is used to train all the models to classify skin cancer photos in a binary manner, and it is shown that the VGG19 outperforms any pre-trained CNN in a comparable manner. Fig. 5 demonstrates the customization for the categorization of skin cancer and the diagram of these pre-trained CNNs.

3.6. DenseNet201 [11,43,66]

A convolution neural network with 201 layers is called DenseNet201 [43]. In DenseNet201, each layer collects information from the layers that came before it. This design makes use of a simple connection scheme to guarantee all available data flow between layers during both backward and forward computing to address the vanishing gradient problem. The layers are interconnected in such a way that each following layer receives input from each by use of its own feature maps, the previous layers. The entire architecture has been broken up into numerous densely connected units to aid in down sampling. Transition layers, which perform pooling and convolution, divide these dense blocks. Four completely connected layers with the activation function "relu" are placed on top of this pre-trained architecture. The last output layer is where classification is performed, and sigmoid is used as the activation function.

4. Result and discussion

The proposed approach, along with every other technique considered for comparison and assessment, was implemented in this section using the Python programming language and the Scikit-Learn package, alongside Keras, a high-level API based on TensorFlow 2.0. To conduct deep learning research, train, and test models, we utilize Google Colab Pro, a cloud-based platform.

4.1. Dataset description

The experiment was conducted using the ISIC 2020 dataset [79], which comprises 33,126 dermoscopic training photos depicting distinct skin lesions from over 2000 patients. These lesions include both benign and malignant cases. Each image is associated with a unique patient identification, ensuring a comprehensive representation of different conditions. Malignant diagnoses were confirmed through biopsy, while benign diagnoses were validated through expert consensus.

For training and validation, 70 % of the dataset was allocated to training, while 30 % was reserved for validation and testing purposes. During training, a batch size of 128 and an epoch count of 50 were utilized to optimize model performance and convergence.

The Memorial Sloan Kettering Cancer Centre, the Hospital Clinic de Barcelona, the Medical School of the University of Athens, The University of Queensland, and the Medical University of Vienna are just a few of the institutions from which images for the collection were compiled. Images of 32,543 benign and 585 malignant moles are included in the dataset. Each image has been reduced in size to a low-resolution RGB file (96 96x3) [26].

Fig. 6 displaying random data samples of melanoma skin cancer images is a valuable step in understanding the nature and quality of the dataset. It offers insights into the diversity of cases, the potential challenges in data preprocessing, and the visual characteristics

that the model needs to learn. Proper dataset curation and quality assurance are essential for building accurate and reliable melanoma detection models, which can have a significant impact on early diagnosis and patient outcomes. Histopathology has validated every malignant diagnosis, while the diagnosis for benign conditions has either been supported by expert consensus, extensive observation, or histopathology. It worth mentioning that class imbalance issue can affect the performance and reliability of machine learning models. Proper handling of class imbalance through various techniques is crucial to ensure that the model is both accurate and effective in detecting malignant skin lesions, ultimately contributing to improved patient outcomes.

The other hyperparameters were set as follows. The gamma is set to be 0.001, weight decay is configured to 0.0005, and learning rate was set to 0.001. Furthermore, epsilon and momentum were adjusted to 0.000001 and 0.99, respectively, in the BN layer. We chose Adam to be the optimizer after several tests comparing the results of SGD, Adam, and RMSProp. The suggested strategy for the categorization of skin cancer images was assessed using the metrics from the challenge evaluation. The area under the Receiver Operating Characteristic curve (AUC), F-measure, specificity, accuracy, and sensitivity are the assessment criteria. We employ Equations (3)–(8), to assess the effectiveness of the proposed system. The mathematical formulation of these measures is outlined as follows. Accuracy reflects the proximity of measurements to a predetermined value, while precision measures the consistency of measurements among themselves. Sensitivity (True Positive rate) quantifies the proportion of true positives correctly identified. Additionally, the F1-score is computed based on the precision and sensitivity of the test, representing the harmonic mean of these two metrics.

$$\mathrm{Acc} = \frac{\mathrm{tp} + \mathrm{tn}}{\mathrm{tp} + \mathrm{fn} + \mathrm{fp} + \mathrm{tn}} \tag{3}$$

$$\mathrm{Pr} = \frac{\mathrm{tp}}{\mathrm{tp} + \mathrm{fp}} \tag{4}$$

$$\mathrm{Sn} = \frac{\mathrm{tp}}{tp + fn} \tag{5}$$

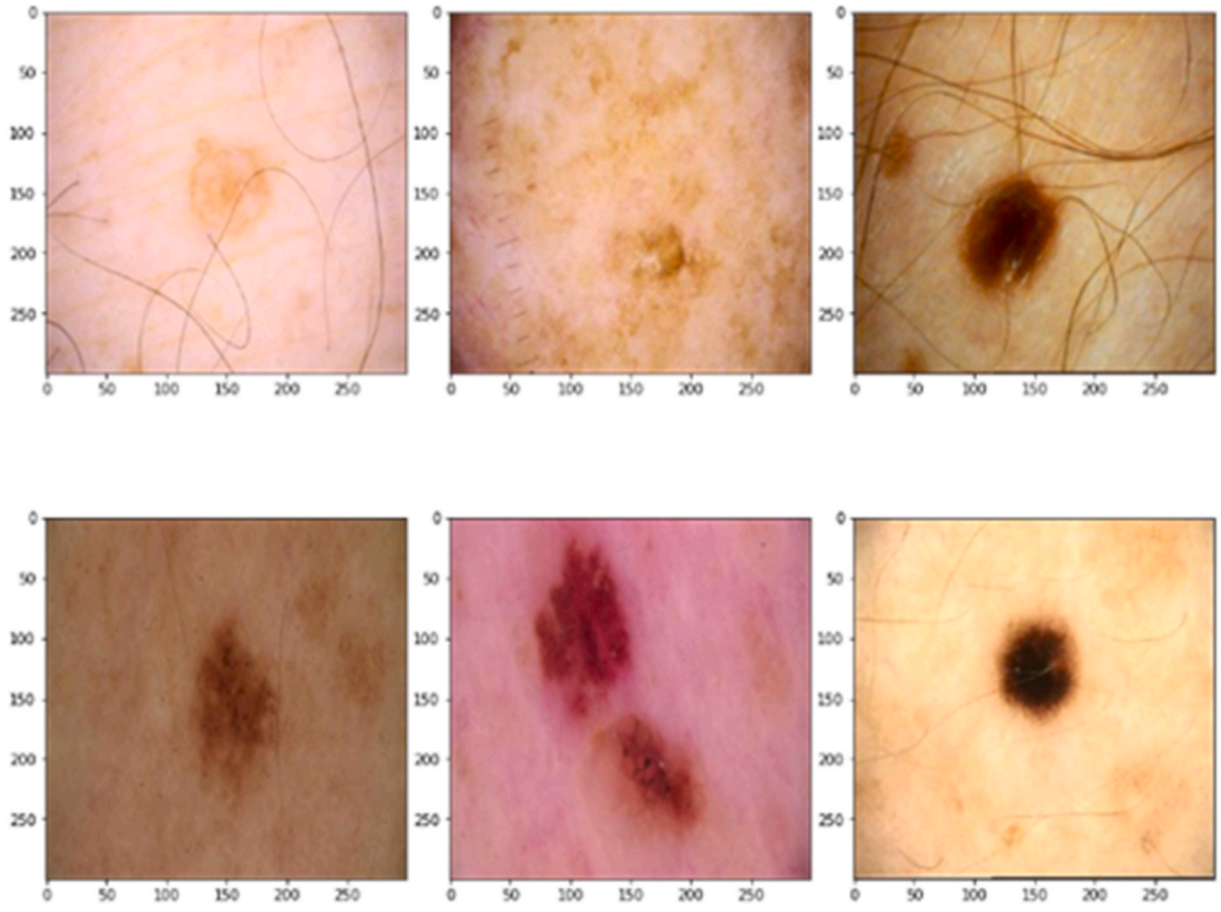


Fig. 6. Random data samples of melanoma skin cancer detection.

$$\mathrm{Sp} = \frac{\mathrm{tn}}{fp + tn} \tag{6}$$

$$\mathrm{F-measure} = \frac{2 \times \mathrm{p} \times \mathrm{r}}{\mathrm{p} + \mathrm{r}} \tag{7}$$

$$\mathrm{G-mean} = \sqrt{sn \times sp} \tag{8}$$

Table 2 compares each of these models based on metrics including accuracy, precision, recall, and F1-score. The DenseNet201 is found to perform better than other models. In this case, we consider every performance indicator needed to assess the suggested model. While some models do well in certain areas, they are unable to produce consistent outcomes across all performance matrices. Resolving the deep model's vanishing gradients issue with the use of skip and residual connections makes ResNet, or the ResNet152v2 variation, the second-best performance behind ResNet50. We tried different optimizers, validation splits, and batch sizes, in addition to the Keras Tuner to do hyperparameter tuning. The most promising set of hyperparameters used in the suggested model. The best set of parameters for the learning process in our model is determined by using the Keras Tuner package. In addition, we conducted experiments using slightly different hypermeters to confirm Keras Tuner's findings. There is also a comparison between the ML classification model with pretrained models. Table 3 displays the testing performance metrics. It is evident from this table that the KNN, DT, Logistic Regression model outperforms the other machine learning models in every performance matrix. Out of all the machine learning models.

Fig. 3 provides layers details of the pretrained network structures, while in-depth details are available in the literature. The performance of the pre-trained models is illustrated in Table 2, ranking the networks according to their accuracy concerning the number of layers. Figs. 7 and 8, presenting F1 scores for pretrained and ML models with respect to benign and malignant classes, The proposed E-VGG16 model acthived 89 % F1 score on benign and 87 % on malignant class. It aids in model selection, threshold optimization, bias detection, and clinical decision-making, ultimately contributing to more accurate and effective diagnostic tools. The skin cancer segmentation shown in Fig. 9 highlights the significance of augmenting data sets for enhancing the dataset and the potential advantages that accrue to a segmentation model. Fig. 10, displaying the confusion matrix for a CNN used in skin cancer classification, is a valuable tool for assessing the model performance in a medical context. It provides essential metrics for evaluating the model's accuracy, precision, recall, and clinical relevance, helping to make informed decisions about its readiness for clinical use and identifying areas for improvement.

For improved model training, generalization, and overall performance, data augmentation can have an impact on clinical practice in the realm of skin cancer diagnosis and treatment. It is critical for ensuring image segmentation models are able to accurately identify skin cancer lesions across different environments. This understanding is based on comprehending the nature and quality of a dataset. It provides insight into variation in cases, data preprocessing problems that may arise and visual attributes which the model must recognize. Dataset curation should be done properly and ensured of its quality so as to come up with melanoma detection models that are accurate and reliable enough to lead to early diagnosis hence better patient outcomes.

We used the Kaggle platform, which runs Linux on an Nvidia K80 GPU and supports Python 3.7, Keras 2.4, and Tensor Flow 2.4, for our experiments. The Kaggle platform was selected because it offers a large library of machine learning and deep learning resources and supports the quicker GPU training of ensemble models or deep learning models. The E-VGG19 model demonstrated an accuracy of

Table 2
Performance Comparison of pre-trained models and ML methods on Skin cancer detection. (B- Benign, M – Malignant).

Methods	Classes	#Precison	#Recall	#F1-Score	#Accuracy
SVM	B	0.79	0.74	0.79	0.79
	M	0.70	0.71	0.70	
Linear SVM	B	0.76	0.70	0.75	0.75
	M	0.75	0.72	0.73	
DT	B	0.74	0.75	0.70	0.74
	M	0.70	0.70	0.69	
KNN	B	0.59	0.65	0.60	0.60
	M	0.70	0.55	0.61	
LR	B	0.80	0.84	0.75	0.75
	M	0.82	0.81	0.73	
VGG19	B	0.83	0.85	0.82	0.82
	M	0.81	0.81	0.83	
ResNet152v2	B	0.81	0.82	0.83	0.82
	M	0.80	0.70	0.82	
InceptionV3	B	0.80	0.88	0.83	0.84
	M	0.80	0.85	0.84	
ResNet50	B	0.80	0.83	0.86	0.86
	M	0.82	0.81	0.80	
DenseNet201	B	0.83	0.83	0.86	0.86
	M	0.78	0.81	0.80	
Proposed E-VGG19	**B**	**0.87**	**0.93**	**0.89**	**0.88**
	M	**0.94**	**0.86**	**0.87**	

Table 3
Comparative performance of VGG19 and E-VGG19 pre-trained networks using different ML classifiers.

Pretrained Model	Classifier	Optimizer	Epoch	Batch Size	Learning Rate	Accuracy	Sensitivity
VGG19	SVM	A	25	32	0.0001	0.83	0.84
		RP	50	64	0.001	0.82	0.82
	Linear SVM	A	25	32	0.0001	0.82	0.82
		RP	50	64	0.001	0.77	0.82
	DT	A	25	32	0.0001	0.79	0.81
		RP	50	64	0.001	0.82	0.81
	KNN	A	25	32	0.0001	0.80	0.83
		RP	50	64	0.001	0.79	0.83
	LR	A	25	32	0.0001	0.78	0.82
		RP	50	64	0.001	0.80	0.83
E-VGG19	SVM	A	25	32	0.0001	0.84	0.82
		RP	50	64	0.001	0.87	0.85
	Linear SVM	A	25	32	0.0001	0.88	0.85
		RP	50	64	0.001	0.87	0.87
	DT	A	25	32	0.0001	0.87	0.87
		RP	50	64	0.001	0.85	0.86
	KNN	A	25	32	0.0001	0.83	0.88
		RP	50	64	0.001	0.82	0.88
	LR	A	25	32	0.0001	0.88	0.87
		RP	50	64	0.001	0.88	0.87

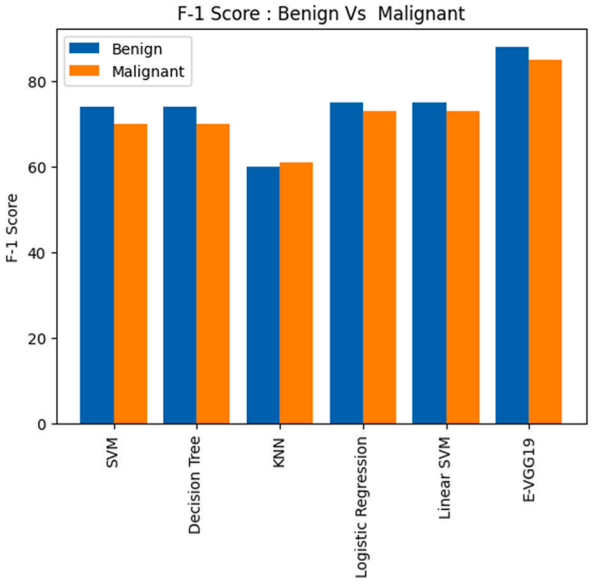


Fig. 7. F1 score comparison with ML classifiers against E-VGG19.

88 % on the training dataset and 91 % on the validation dataset, according to our observations. The model undergoes 50 iterations of training, when using the Adam optimizer, and the loss function is binary cross entropy. Furthermore, additional deep learning models, including InceptionResNetV2, DenseNet201, ResNet50, and InceptionV3, are also being explored and experimented with for the purpose of detecting skin cancer.

Table 2 compares each of these models based on metrics including accuracy, precision, recall, and F1-score. The E-VGG19 is found to perform better than other models. In this case, we consider every performance indicator needed to assess the suggested model. While some models do well in certain areas, they are unable to produce consistent outcomes across all performance matrices. Resolving the deep model's vanishing gradients issue with the use of skip and residual connections makes ResNet, or the ResNet152v2 variation, the second-best performance behind VGG19. We tried different optimizers, validation splits, and batch sizes, in addition to the Keras Tuner to do hyperparameter tuning. The most optimal collection of parameters for the learning method is determined by using the Keras Tuner package. In addition, we conducted other experiments using slightly different hypermeters to confirm Keras Tuner's findings.

The following confusion matrix reported in Fig. 10 is used to determine the classification's assessment criteria for the main DL pretrained models. On one hand, True Positive (TP) refers to the instances that were correctly classified as positive by the model where True Negative (TN) signifies instances that were correctly classified as negative by the model. On the other hand, False Positive (FP) indicates the instances that were falsely classified as positive by the model. Also, False Negative (FN) represents the number of

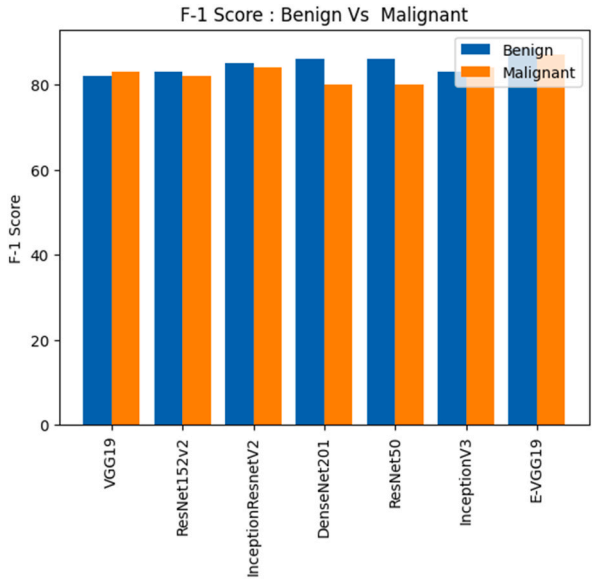


Fig. 8. F1 score comparison with DL pretrained models against E-VGG19.

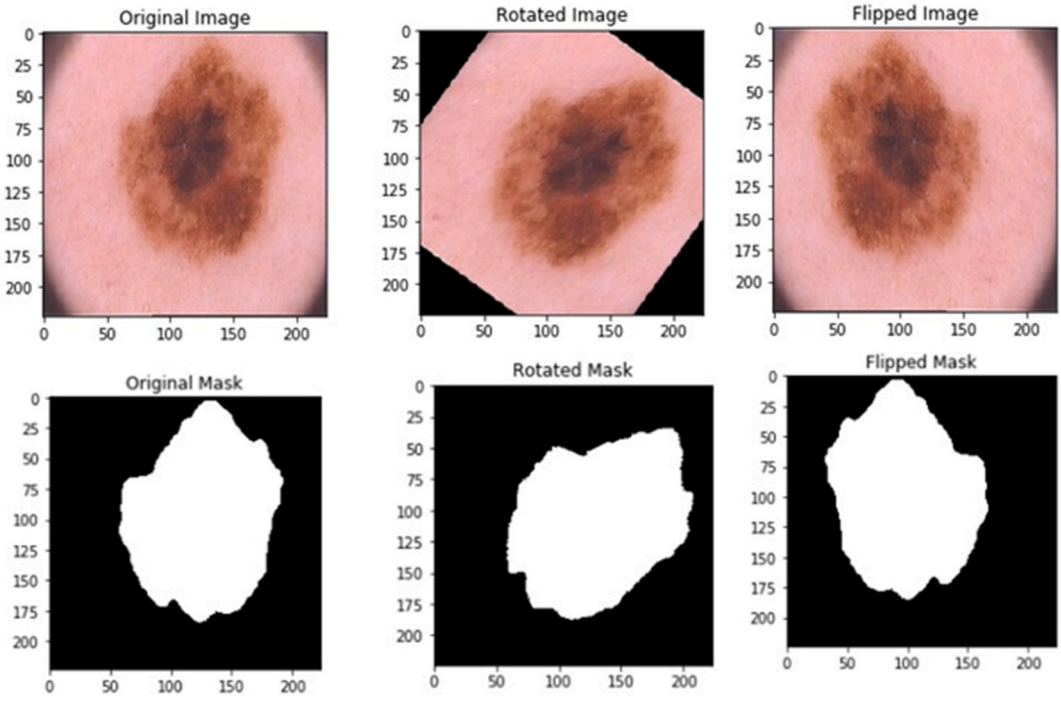


Fig. 9. Skin Cancer Segmentation with data Augmentation techniques.

incorrect predictions where the model predicted the class as negative (False) when the actual class was positive (True).

The Receiver Operating Characteristic (ROC) curve combines these measurements to create a graphical representation. This curve shows how many times a model makes classification mistakes and FN or FP rates on top of that. Moreover, ROC curve can also be used in creating AUC plot. Specifically speaking here, AUC follows the ROC as probability curve which represent the degree of separability for any given test result. The ROC curve for the benign-vs-malignant lesions is shown in Fig. 11. The classifier gives a TPR of 0.89, a FPR of 0.20 and a FNR of 0.11 for the malignant lesions at the optimal cut-off value of 0.42. The area under the ROC curve is 0.94 for the benign-versus-malignant classification problem. A model's ability to do classification work increases with its AUC. A model that has a near-one AUC is considered optimal and has high separability metrics. Conversely, an AUC close to 0 indicates a poor separability

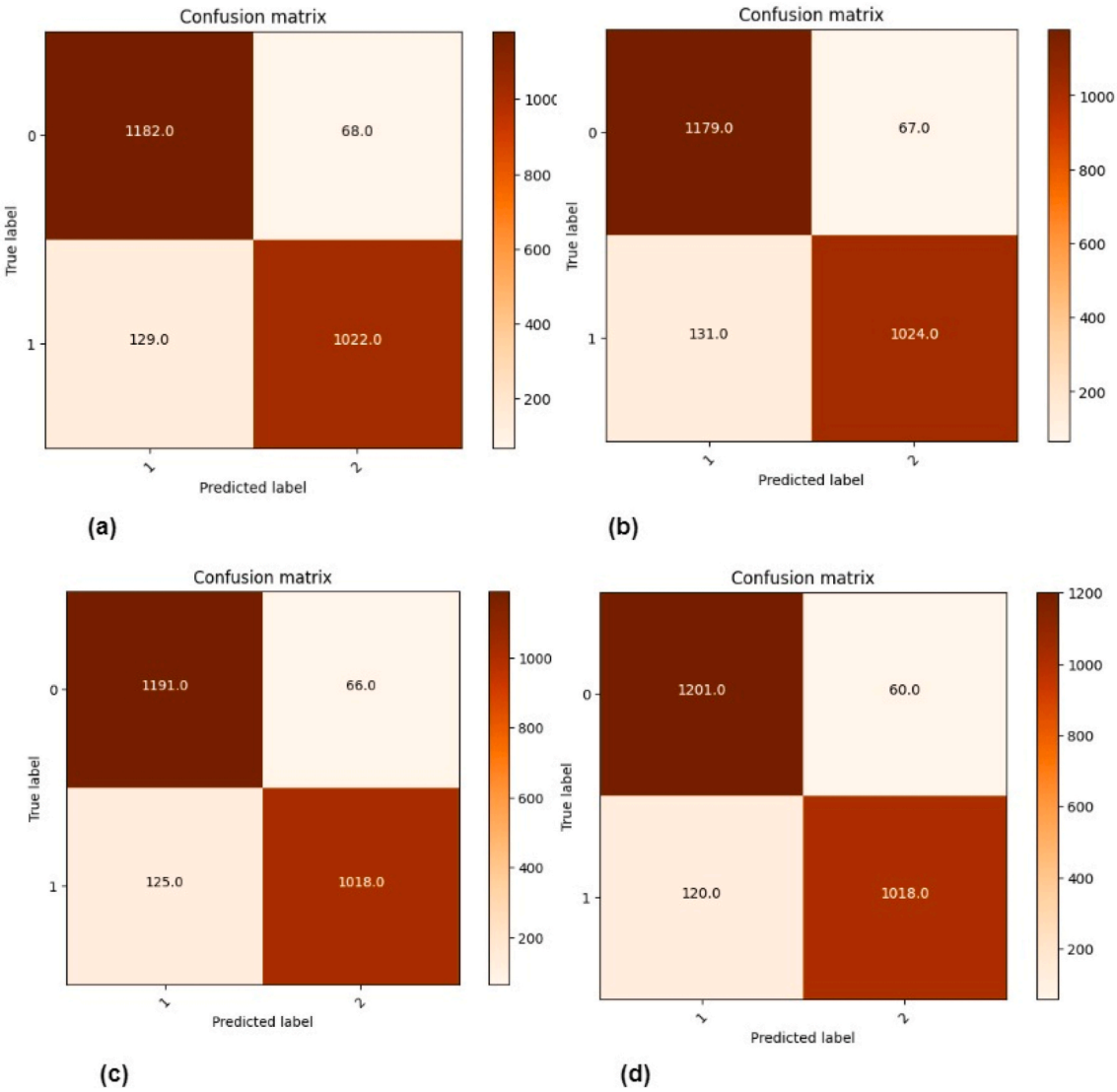


Fig. 10. Classification of skin cancer lesion using pretrained models (a) InceptionV3, (b) ResNet50 and (c) DenseNet201, and d) ResNet152v2.

metric. The probabilities of TP and FP represent sensitivity and specificity, respectively, the estimated AUC may be expressed in the following manner.

Fig. 12 presents a visual depiction of the precision and decrease across epochs for the two models utilized in our framework. To attain optimal performance, we adopted the modified DenseNet201 architecture with pre-trained ImageNet weights and subsequently fine-tuned the fully connected layers. In Fig. 12 (a), the training loss achieved a remarkable 0.17 %, while the validation loss stood at 0.18 %. Additionally, this model demonstrated a superior training accuracy of 91.01 % and a commendable validation accuracy of 84.22 %. The findings show how effective our method is and the high-quality performance that can be achieved by using ResNet152v2.

5. Conclusion

This research study presents novel (E-VGG16) model by using machine learning and transfer learning techniques for real-time artificial intelligence-based skin cancer detection. In this paper, we have combine transfer learning (pretrained)and machine learning classifers to come up with an efficient skin cancer classifier. We used transfer-learning which enabled us take advantage of feature extraction capability in pre-trained DL models, greatly reducing the need for manual features creation as well as making our system more robust. The outcomes show that these combinations of models were able to correctly diagnose skin lesions particularly E-VGG19. In turn, application of transfer learning together with good fine-tuning technique and data augmentation has led to significant improvement in recognition accuracy. These findings would totally transform the future diagnosis of skin cancer since they have great implications for real time skin cancer diagnoses. Using a mixture of deep learning and machine concepts we were able to create a strong

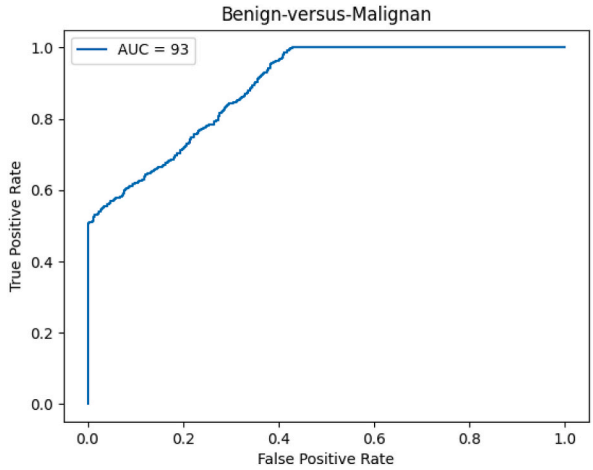


Fig. 11. ROC curve of benign-versus-malignant.

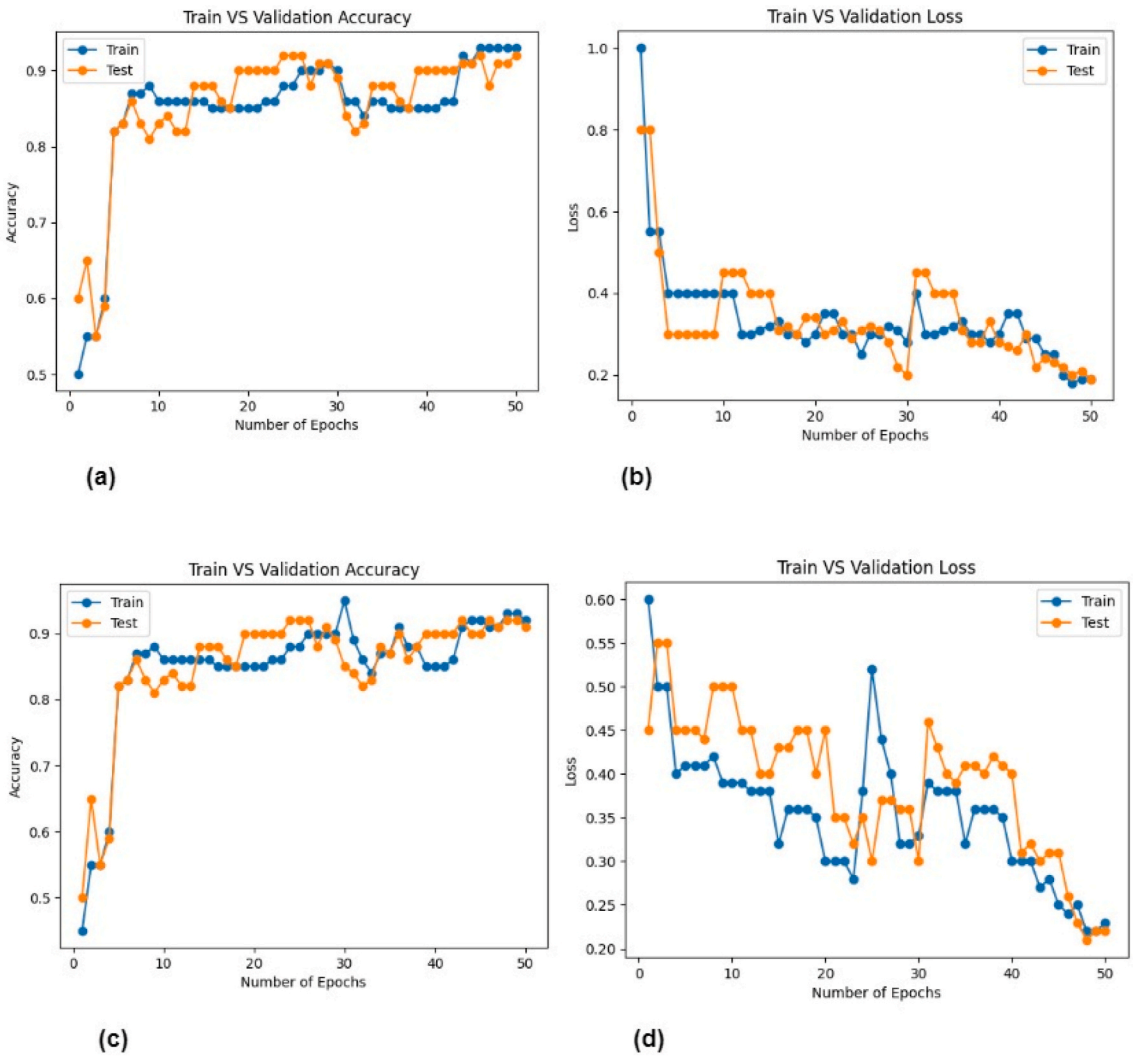


Fig. 12. Training and Validation accuracy ResNet50 (a), (b) and a) ResNet152v2 (c) & (d).

tool which can help medical professionals accurately identify skin cancers more instantly than ever before. Given our study, we recommend further investigation on these methods over wider range of subjects. Furthermore, the integration of real-time capabilities into the system to enable instantaneous skin cancer detection presents an exciting avenue for future research, with the potential to make a substantial impact on healthcare, particularly in the field of dermatology. Researchers may use sectional classification in the future to examine techniques for detecting skin cancer. There are other categories within traditional segmentation, such as threshold-based, region-based, and clustering-based. In a similar vein, colour, shape, and texture features should be distinguished when classifying the extracted features. While CNN continues to produce excellent results, there are still certain constraints, such as those related to generalization and optimization, which may be resolved in the future with the use of quantum computing.

CRediT authorship contribution statement

Irfan Ali Kandhro: Software, Visualization, Writing – original draft. **Selvakumar Manickam:** Software, Visualization, Writing – original draft. **Kanwal Fatima:** Investigation, Software, Visualization. **Mueen Uddin:** Conceptualization, Methodology, Writing – review & editing. **Urooj Malik:** Investigation, Software. **Anum Naz:** Investigation, Software, Visualization. **Abdulhalim Dandoush:** Conceptualization, Methodology, Supervision, Validation, Writing – review & editing.

Declaration of competing interest

As the corresponding author and on behalf of all co-authors, I declare that we have no financial or personal relationships with other individuals or organizations that could inappropriately influence (bias) our work presented in the manuscript titled "Performance Evaluation of E-VGG19 Model: Enhancing Real-Time Skin Cancer Detection and Classification." We affirm that we have no conflicts of interest, including but not limited to employment, consultancies, stock ownership, honoraria, paid expert testimony, patent applications/registrations, and grants or other funding that could influence the integrity or objectivity of our research.

Acknowledgments

Open Access funding provided by the Qatar National Library. Private comment to editors: This follows the Agreement between Qatar and Elsevier, more details can be found here: https://www.elsevier.com/open-access/agreements/qatar.

References

[1] I. de Vere Hunt, J. Lester, E. Linos, Insufficient evidence for screening reinforces need for primary prevention of skin cancer, JAMA Intern. Med. 183 (6) (2023) 509–511.
[2] R. Javed, M.S.M. Rahim, T. Saba, A. Rehman, A comparative study of features selection for skin lesion detection from dermoscopic images, Network Modeling Analysis in Health Informatics and Bioinformatics 9 (2020) 1–13.
[3] H. Tabrizchi, S. Parvizpour, J. Razmara, An improved VGG model for skin cancer detection, Neural Process. Lett. 55 (4) (2023) 3715–3732.
[4] S.H. Kassani, P.H. Kassani, A comparative study of deep learning architectures on melanoma detection, Tissue Cell 58 (2019) 76–83.
[5] A. Adegun, S. Viriri, Deep learning techniques for skin lesion analysis and melanoma cancer detection: a survey of state-of-the-art, Artif. Intell. Rev. 54 (2021) 811–841.
[6] H. Ganster, P. Pinz, R. Rohrer, E. Wildling, M. Binder, H. Kittler, Automated melanoma recognition, IEEE Trans. Med. Imag. 20 (3) (2001) 233–239.
[7] M.R. Hasan, M.I. Fatemi, M. Monirujjaman Khan, M. Kaur, A. Zaguia, Comparative analysis of skin cancer (benign vs. malignant) detection using convolutional neural networks, Journal of Healthcare Engineering 2021 (2021).
[8] H. Bhatt, V. Shah, K. Shah, R. Shah, M. Shah, State-of-the-art machine learning techniques for melanoma skin cancer detection and classification: a comprehensive review, Intelligent Medicine 3 (3) (2023) 180–190.
[9] Maryam Naqvi, et al., Skin cancer detection using deep learning—a review, Diagnostics 13.11 (2023) 1911.
[10] A.M. Smak Gregoor, T.E. Sangers, L.J. Bakker, L. Hollestein, C.A. Uyl–de Groot, T. Nijsten, M. Wakkee, An artificial intelligence based app for skin cancer detection evaluated in a population based setting, NPJ digital medicine 6 (1) (2023) 90.
[11] J.V. Tembhurne, N. Hebbar, H.Y. Patil, T. Diwan, Skin cancer detection using ensemble of machine learning and deep learning techniques, Multimed. Tool. Appl. (2023) 1–24.
[12] M. Obayya, M.A. Arasi, N.S. Almalki, S.S. Alotaibi, M. Al Sadig, A. Sayed, Internet of things-assisted smart skin cancer detection using metaheuristics with deep learning model, Cancers 15 (20) (2023) 5016.
[13] N.A. Al-Dmour, M. Salahat, H.K. Nair, N. Kanwal, M. Saleem, N. Aziz, Intelligence skin cancer detection using IoT with a fuzzy expert system, in: 2022 International Conference on Cyber Resilience (ICCR), IEEE, Dubai, United Arab Emirates, 2022, October, pp. 1–6.
[14] Maryam Tahir, et al., DSCC_Net: multi-classification deep learning models for diagnosing of skin cancer using dermoscopic images, Cancers 15 (7) (2023) 2179.
[15] O.T. Jones, R.N. Matin, M. van der Schaar, K.P. Bhayankaram, C.K.I. Ranmuthu, M.S. Islam, F.M. Walter, Artificial intelligence and machine learning algorithms for early detection of skin cancer in community and primary care settings: a systematic review, The Lancet Digital Health 4 (6) (2022) e466–e476.
[16] Andre GC. Pacheco, Renato A. Krohling, The impact of patient clinical information on automated skin cancer detection, Comput. Biol. Med. 116 (2020) 103545.
[17] M. Toğaçar, Z. Cömert, B. Ergen, Intelligent skin cancer detection applying autoencoder, MobileNetV2 and spiking neural networks, Chaos, Solit. Fractals 144 (2021) 110714.
[18] M. Hasan, S.D. Barman, S. Islam, A.W. Reza, Skin cancer detection using convolutional neural network, in: Proceedings of the 2019 5th International Conference on Computing and Artificial Intelligence, Bali, Indonesia, 2019, April, pp. 254–258.
[19] B. Sreedhar, M.S. Be, M.S. Kumar, A comparative study of melanoma skin cancer detection in traditional and current image processing techniques, in: 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) at SCAD Institute of Technology, IEEE, Palladam, India, 2020, October, pp. 654–658.
[20] M. Dildar, S. Akram, M. Irfan, H.U. Khan, M. Ramzan, A.R. Mahmood, M.H. Mahnashi, Skin cancer detection: a review using deep learning techniques, Int. J. Environ. Res. Publ. Health 18 (10) (2021) 5479.
[21] Duggani Keerthana, et al., Hybrid convolutional neural networks with SVM classifier for classification of skin cancer, in: Biomedical Engineering Advances, vol. 5, 2023 100069.
[22] N. Alfed, F. Khelifi, A. Bouridane, Improving a bag of words approach for skin cancer detection in dermoscopic images, in: 2016 International Conference on Control, Decision and Information Technologies (CoDIT), IEEE, Saint Julian, Malta, 2016, April, pp. 24–27.

[23] N. Durgarao, G. Sudhavani, Diagnosing skin cancer via C-means segmentation with enhanced fuzzy optimization, IET Image Process. 15 (10) (2021) 2266–2280.
[24] M. Vidya, M.V. Karki, Skin cancer detection using machine learning techniques, in: 2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), IEEE, GITAM Bangalore, India, 2020, July, pp. 1–5.
[25] J.C. Kavitha, A. Suruliandi, Texture and color feature extraction for classification of melanoma using SVM, in: 2016 International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE'16), IEEE, 2016, January, pp. 1–6.
[26] J.C. Kavitha, A. Suruliandi, D. Nagarajan, T. Nadu, Melanoma detection in dermoscopic images using global and local feature extraction, International Journal of Multimedia and Ubiquitous Engineering 12 (5) (2017) 19–28.
[27] H.S.K.A. Putri, C.A. Sari, E.H. Rachmawanto, Classification of skin diseases types using naïve Bayes classifier based on local binary pattern features, in: 2020 International Seminar on Application for Technology of Information and Communication (iSemantic), IEEE, 2020, September, pp. 61–66.
[28] N. Alfed, F. Khelifi, Bagged textural and color features for melanoma skin cancer detection in dermoscopic and standard images, Expert Syst. Appl. 90 (2017) 101–110.
[29] F.K. Nezhadian, S. Rashidi, Melanoma skin cancer detection using color and new texture features, in: 2017 Artificial Intelligence and Signal Processing Conference (AISP), Shiraz University, Iran, 2017, October, pp. 1–5. IEEE.
[30] Y. Filali, A. Ennouni, M.A. Sabri, A. Aarab, A study of lesion skin segmentation, features selection and classification approaches, in: 2018 International Conference on Intelligent Systems and Computer Vision (ISCV), IEEE, Fez, Morocco, 2018, April, pp. 1–7.
[31] S.S. Zareen, S. Guangmin, Y. Li, M. Kundi, S. Qadri, S.F. Qadri, A.H. Khan, A machine vision approach for classification of skin cancer using hybrid texture features, Comput. Intell. Neurosci. 2022 (2022).
[32] K. Melbin, Y.J.V. Raj, Integration of modified ABCD features and support vector machine for skin lesion types classification, Multimed. Tool. Appl. 80 (6) (2021) 8909–8929.
[33] R. Gaonkar, K. Singh, G.R. Prashanth, V. Kuppili, Lesion analysis towards melanoma detection using soft computing techniques, Clinical Epidemiology and Global Health 8 (2) (2020) 501–508.
[34] F. Afza, M.A. Khan, M. Sharif, T. Saba, A. Rehman, M.Y. Javed, Skin lesion classification: an optimized framework of optimal color features selection, in: 2020 2nd International Conference on Computer and Information Sciences (ICCIS), IEEE, Jouf University, Aljouf, Kingdom of Saudi Arabia, 2020, October, pp. 1–6.
[35] G. Arora, A.K. Dubey, Z.A. Jaffery, A. Rocha, Bag of feature and support vector machine based early diagnosis of skin cancer, Neural Comput. Appl. (2022) 1–8.
[36] I. ul haq, J. Amin, M. Sharif, M. Almas Anjum, Skin lesion detection using recent machine learning approaches, in: Prognostic Models in Healthcare: AI and Statistical Approaches, Springer Nature Singapore, Singapore, 2022, pp. 193–211.
[37] P. Banasode, M. Patil, N. Ammanagi, A melanoma skin cancer detection using machine learning technique: support vector machine, in: IOP Conference Series: Materials Science and Engineering, IOP Publishing, United Kingdom, 2021, February, 1065, No. 1, p. 012039.
[38] P. Shahi, S. Yadav, N. Singh, N.P. Singh, Melanoma skin cancer detection using various classifiers, in: 2018 5th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON) Gorakhpur, Uttar Pradesh, India, IEEE, 2018, November, pp. 1–5.
[39] C.R. Dhivyaa, K. Sangeetha, M. Balamurugan, S. Amaran, T. Vetriselvi, P. Johnpaul, Skin lesion classification using decision trees and random forest algorithms, J. Ambient Intell. Hum. Comput. (2020) 1–13.
[40] M.A. Arasi, E.S.M. El-Horbaty, A. El-Sayed, Classification of dermoscopy images using naive bayesian and decision tree techniques, in: 2018 1st Annual International Conference on Information and Sciences (AiCIS), IEEE, Fallujah, Iraq, 2018, November, pp. 7–12.
[41] N. Hameed, A. Shabut, M.A. Hossain, A Computer-aided diagnosis system for classifying prominent skin lesions using machine learning, in: 2018 10th Computer Science and Electronic Engineering (CEEC), 2018, pp. 186–191.
[42] V.R. Balaji, S.T. Suganthi, R. Rajadevi, V.K. Kumar, B.S. Balaji, S. Pandiyan, Skin disease detection and segmentation using dynamic graph cut algorithm and classification through Naive Bayes classifier, Measurement 163 (2020) 107922.
[43] W. Abbes, D. Sellami, S. Marc-Zwecker, C. Zanni-Merk, Fuzzy decision ontology for melanoma diagnosis using KNN classifier, Multimed. Tool. Appl. 80 (2021) 25517–25538.
[44] H.D. Praveena, K. Sudha, P. Geetha, Support vector machine based melanoma skin cancer detection, J. Univ. Shanghai Sci. Technol. 22 (2020) 1075–1081.
[45] C.L. Chowdhary, M. Alazab, A. Chaudhary, S. Hakak, T.R. Gadekallu, Computer Vision and Recognition Systems Using Machine and Deep Learning Approaches: Fundamentals, Technologies and Applications, Institution of Engineering and Technology, 2021.
[46] C.L. Chowdhary, G.T. Reddy, B.D. Parameshachari, Computer Vision and Recognition Systems: Research Innovations and Trends, CRC Press, 2022.
[47] M. Elgamal, Automatic skin cancer images classification, Int. J. Adv. Comput. Sci. Appl. 4 (3) (2013).
[48] A. Sáez, J. Sánchez-Monedero, P.A. Gutiérrez, C. Hervás-Martínez, Machine learning methods for binary and multiclass classification of melanoma thickness from dermoscopic images, IEEE Trans. Med. Imag. 35 (4) (2015) 1036–1045.
[49] H. Alquran, I.A. Qasmieh, A.M. Alqudah, S. Alhammouri, E. Alawneh, A. Abughazaleh, F. Hasayen, The melanoma skin cancer detection and classification using support vector machine, in: 2017 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), IEEE Aqaba, Jordan, 2017, October, pp. 1–5.
[50] M.A. Ismail, N. Hameed, J. Clos, Deep learning-based algorithm for skin cancer classification, in: Proceedings of International Conference on Trends in Computational and Cognitive Engineering: Proceedings of TCCE 2020, Springer, Singapore, 2021, pp. 709–719.
[51] J. Zhang, Y. Xie, Y. Xia, C. Shen, Attention residual learning for skin lesion classification, IEEE Trans. Med. Imag. 38 (9) (2019) 2092–2103.
[52] S.S. Chaturvedi, J.V. Tembhurne, T. Diwan, A multi-class skin Cancer classification using deep convolutional neural networks, Multimed. Tool. Appl. 79 (39–40) (2020) 28477–28498.
[53] D.N. Dorrell, L.C. Strowd, Skin cancer detection technology, Dermatol. Clin. 37 (4) (2019) 527–536.
[54] Tehseen Mazhar, et al., The role of machine learning and deep learning approaches for the detection of skin cancer, Healthcare 11 (3) (2023). MDPI.
[55] J. Verstockt, S. Verspeek, F. Thiessen, W.A. Tjalma, L. Brochez, G. Steenackers, Skin cancer detection using infrared thermography: measurement setup, procedure and equipment, Sensors 22 (9) (2022) 3327.
[56] H. Nahata, S.P. Singh, Deep learning solutions for skin cancer detection and diagnosis. Machine Learning with Health Care Perspective: Machine Learning and Healthcare, 2020, pp. 159–182.
[57] N. Priyadharshini, N. Selvanathan, B. Hemalatha, C. Sureshkumar, A novel hybrid Extreme Learning Machine and Teaching–Learning-Based Optimization algorithm for skin cancer detection, Healthcare Analytics 3 (2023) 100161.
[58] E.G. Dobre, M. Surcel, C. Constantin, M.A. Ilie, A. Caruntu, C. Caruntu, M. Neagu, Skin cancer pathobiology at a glance: a focus on imaging techniques and their potential for improved diagnosis and surveillance in clinical cohorts, Int. J. Mol. Sci. 24 (2) (2023) 1079.
[59] N. ul Huda, R. Amin, S.I. Gillani, M. Hussain, A. Ahmed, H. Aldabbas, Skin cancer malignancy classification and segmentation using machine learning algorithms, J. Miner. Met. Mater. Soc. (2023) 1–15.
[60] S. Maurya, S. Tiwari, M.C. Mothukuri, C.M. Tangeda, R.N.S. Nandigam, D.C. Addagiri, A review on recent developments in cancer detection using Machine Learning and Deep Learning models, Biomed. Signal Process Control 80 (2023) 104398.
[61] M. Hossain, K. Sadik, M.M. Rahman, F. Ahmed, M.N.H. Bhuiyan, M.M. Khan, Convolutional neural network based skin cancer detection (malignant vs benign), in: 2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), IEEE, Vancouver, Canada, 2021, October, pp. 141–147.
[62] L. Yu, H. Chen, Q. Dou, J. Qin, P.A. Heng, Automated melanoma recognition in dermoscopy images via very deep residual networks, IEEE Trans. Med. Imag. 36 (4) (2016) 994–1004.
[63] T. DeVries, D. Ramachandram, Skin Lesion Classification Using Deep Multi-Scale Convolutional Neural Networks, 2017 arXiv preprint arXiv:1703.01402.
[64] A. Mahbod, G. Schaefer, C. Wang, R. Ecker, I. Ellinge, Skin lesion classification using hybrid deep neural networks, in: ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, Brighton, United Kingdom, 2019, May, pp. 1229–1233.

[65] A. Namozov, D. Ergashev, Y. Im Cho, Adaptive activation functions for skin lesion classification using deep neural networks, in: 2018 Joint 10th International Conference on Soft Computing and Intelligent Systems (SCIS) and 19th International Symposium on Advanced Intelligent Systems (ISIS), IEEE ,Toyama International Conference Center, Japan, 2018, December, pp. 232–235.
[66] A. Rezvantalab, H. Safigholi, S. Karimijeshni, Dermatologist level dermoscopy skin cancer classification using different deep learning convolutional neural networks algorithms, arXiv preprint arXiv:1810.10348 (2018).
[67] M.A.A. Milton, Automated Skin Lesion Classification Using Ensemble of Deep Neural Networks in Isic 2018: Skin Lesion Analysis towards Melanoma Detection Challenge, arXiv preprint arXiv:1901.10802, 2019.
[68] C.J. Ho, M. Calderon-Delgado, C.C. Chan, M.Y. Lin, J.W. Tjiu, S.L. Huang, H.H. Chen, Detecting mouse squamous cell carcinoma from submicron full-field optical coherence tomography images by deep learning, J. Biophot. 14 (1) (2021) e202000271.
[69] J. Kawahara, A. BenTaieb, G. Hamarneh, Deep features to classify skin lesions, in: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), IEEE, 2016, April, pp. 1397–1400.
[70] N. Nida, A. Irtaza, A. Javed, M.H. Yousaf, M.T. Mahmood, Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy C-means clustering, Int. J. Med. Inf. 124 (2019) 37–48.
[71] T. Yoshida, M.E. Celebi, G. Schaefer, H. Iyatomi, Simple and effective pre-processing for automated melanoma discrimination based on cytological findings, in: 2016 IEEE International Conference on Big Data (Big Data), IEEE, Washington D.C., USA, 2016, December, pp. 3439–3442.
[72] H. Liao, Y. Li, J. Luo, Skin disease classification versus skin lesion characterization: achieving robust diagnosis using multi-label deep neural networks, in: 2016 23rd International Conference on Pattern Recognition (ICPR), IEEE, Cancun, Mexico, 2016, December, pp. 355–360.
[73] M.S.S. Mahecha, O.J.S. Parra, J.B. Velandia, Design of a system for melanoma detection through the processing of clinical images using artificial neural networks, in: Challenges and Opportunities in the Digital Era: 17th IFIP WG 6.11 Conference on E-Business, E-Services, and E-Society, I3E 2018, Kuwait City, Kuwait, October 30–November 1, 2018, Proceedings, vol. 17, Springer International Publishing, Hawally, Kuwait, 2018, pp. 605–616.
[74] M. Attia, M. Hossny, S. Nahavandi, A. Yazdabadi, Skin melanoma segmentation using recurrent and convolutional neural networks, in: 2017 IEEE 14th International Symposium on Biomedical Imaging, IEEE, 2017, April, pp. 292–296. ISBI 2017).
[75] M.A. Albahar, Skin lesion classification using convolutional neural network with novel regularizer, IEEE Access 7 (2019) 38306–38313.
[76] V. Singh, I. Nwogu, Analyzing skin lesions in dermoscopy images using convolutional neural networks, in: 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE, Kovilpatti, India, 2018, October, pp. 4035–4040.
[77] B.C. Furriel, B.D. Oliveira, R. Prôa, J.Q. Paiva, R.M. Loureiro, W.P. Calixto, M. Giavina-Bianchi, Artificial intelligence for skin cancer detection and classification for clinical environment: a systematic review, Front. Med. 10 (2024) 1305954.
[78] M. Shariff, P. Singh Sm, S. Dp, V. Mh, A.S. Poornima, Convolutional neural network for detection of oral cavity leading to oral cancer from photographic images, International Journal of Computing and Digital Systems 15 (1) (2024) 865–877.
[79] V. Rotemberg, N. Kurtansky, B. Betz-Stablein, L. Caffery, E. Chousakos, N. Codella, M. Combalia, S. Dusza, P. Guitera, D. Gutman, A. Halpern, B. Helba, H. Kittler, K. Kose, S. Langer, K. Lioprys, J. Malvehy, S. Musthaq, J. Nanda, O. Reiter, G. Shih, A. Stratigos, P. Tschandl, J. Weber, P. Soyer, A patient-centric dataset of images and metadata for identifying melanomas using clinical context, Sci. Data 8 (2021) 34, https://doi.org/10.1038/s41597-021-00815-z.