Skin cancer detection using ensemble of machine learning and deep learning techniques

Jitendra V. Tembhurne[1] · **Nachiketa Hebbar**[2] · **Hemprasad Y. Patil**[2] · **Tausif Diwan**[1]

Abstract

Skin cancer is one of the most common forms of cancer, which makes it pertinent to be able to diagnose it accurately. In particular, melanoma is a form of skin cancer that is fatal and accounts for 6 of every 7-skin cancer related death. Moreover, in hospitals where dermatologists have to diagnose multiple cases of skin cancer, there are high possibilities of false negatives in diagnosis. To avoid such incidents, there here has been exhaustive research conducted by the research community all over the world to build highly accurate automated tools for skin cancer detection. In this paper, we introduce a novel approach of combining machine learning and deep learning techniques to solve the problem of skin cancer detection. The deep learning model uses state-of-the-art neural networks to extract features from images whereas the machine learning model processes image features which are obtained after performing the techniques such as Contourlet Transform and Local Binary Pattern Histogram. Meaningful feature extraction is crucial for any image classification roblem. As a result, by combining the manual and automated features, our designed model achieves a higher accuracy of 93% with an individual recall score of 99.7% and 86% for the benign and malignant forms of cancer, respectively. We benchmarked the model on publicly available Kaggle dataset containing processed images from ISIC Archive dataset. The proposed ensemble outperforms both expert dermatologists as well as other state-of-the-art deep

✉ Jitendra V. Tembhurne
  jtembhurne@iiitn.ac.in

  Nachiketa Hebbar
  nachiketa.hebbar2017@vitstudent.ac.in

  Hemprasad Y. Patil
  hemprasad.patil@vit.ac.in

  Tausif Diwan
  tdiwan@iiitn.ac.in

[1] Department of Computer Science & Engineering, Indian Institute of Information Technology, Nagpur, India

[2] School of Electronics Engineering (SENSE), VIT Vellore University, Vellore, India

learning and machine learning methods. Thus, this novel method can be of high assistance to dermatologists to help prevent any misdiagnosis.

Keywords Melanoma · Skin cancer · Deep learning · Machine learning · Contourlet transform

1 Introduction

Melanoma is a fatal form of skin cancer which is often undiagnosed or misdiagnosed as a benign skin lesion. According to the Surveillance Epidemiology and End Results data, melanoma is the sixth most common fatal malignancy in the United States, responsible for 4% of all cancer deaths and 6 of every 7-skin cancer-related deaths [36]. In 2021, the fatality of melanoma is only expected to rise as presented in [2]. The American Cancer Society's estimates for melanoma in the United States for 2021 are about 1,06,110 new melanomas will be diagnosed (about 62,260 in men and 43,850 in women) and about 7180 people are expected to die of melanoma (about 4600 men and 2580 women).

In India specifically, melanoma form of skin is more severe in patients than non-melanoma form of skin cancer [29]. Moreover, percent of melanoma of skin out of all cancers is found highest in Northern region of India for both males and females cited as 1.06 and 0.8 respectively. As per worldwide cases, the percent of melanoma of the skin out of all cancers for both males and females were identified highest in European region which is 48.3 and 53.5, respectively.

For prevention and successful treatment, early detection of skin cancer is essential. Thus, the lives of melanoma patients are totally depended on an accurate and early diagnosis followed by adequate treatment. Due to the critical nature of diagnosis, a high proportion of benign pigmented lesion end up being referred from primary care to specialist care [45]. This leads to an overburdening of scarce medical resources and dermatologists, especially in places where there is shortage of highly trained medical specialists and dermatologists. Physicians often rely on personal experience and evaluate each patient's lesions on a case-to-case basis by taking into account the patient's local lesion patterns in comparison to the entire body.

Without computer-based assistance [5], the respective sensitivity and specificity of lesion classification by the dermatologists were 67.2% (95% confidence interval [CI]: 62.6%–71.7%) and 62.2% (95% CI: 57.6%–66.9%). However, the visual differences between melanoma and benign skin lesions can be very subtle (Fig. 1), making it difficult to distinguish the two cases, even for trained medical experts. A number of automated skin lesion classification techniques are presented in the recent past, as highlighted in the related works section of this paper. However, each algorithm suffers from several drawbacks, primarily because of the problems associated with medical imaging [43]. Due to the lack of large datasets, deep convolution neural networks fail to give very high accuracies, unlike classification using ImageNet dataset where millions of images are available for training. Additionally, as shown in Fig. 1, skin lesions occupy a small section of the dermoscopic image, making it easy for the other parts like skin tissues to interfere with the classification of skin lesions.

For the aforementioned reasons, an intelligent medical imaging-based skin lesion diagnosis system can be a welcome tool to assist a physician in classifying skin lesions. In this work, we are interested in a specific two-class classification problem namely, to determine whether a dermoscopic image containing a skin lesion comprises a melanoma or benign. The proposed
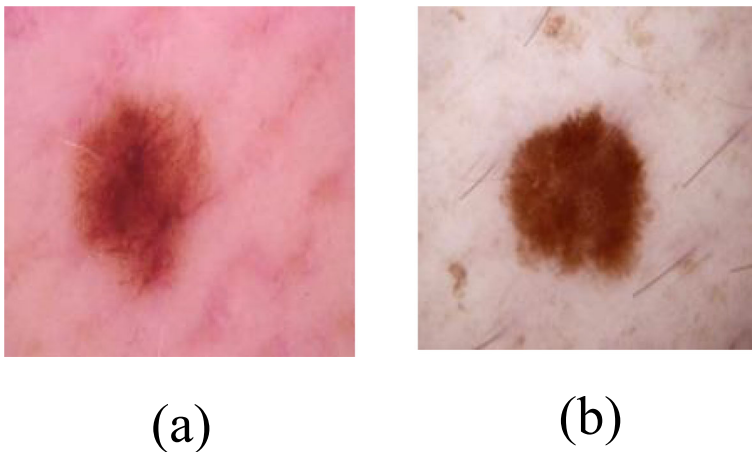
<p align="center">(a)   (b)</p>

Fig. 1 **a** Benign Skin Cancer and **b** Malignant Skin Cancer

framework seeks to build an accurate classifier and minimize the drawbacks of the existing methods.

Prior to this work, either machine learning model such as SVM or logistic regression or deep learning model i.e. CNN were employed for the binary classification of skin cancer [25]. Moreover, the image pre-processing techniques are quite effective before providing features to the ML/DL models. It is evident from the literature that SVM works well as a classifier rather than features' extractor and CNN works well as a features' extractor rather that classifier. Moreover, both of these approaches are based on the effective image pre-processing techniques that were adopted before providing input to the aforementioned ML/DL models. This was the prime motivation behind proposing our ensemble classifier using ML and DL.

In this work, we employ the benefits of Machine Learning (ML) and Deep Learning (DL) for the design of new skin cancer detection model. Machine learning is a subset of Artificial Intelligence (AI) that makes use of mathematical algorithms to learn and understand patterns in the data. To perform well, ML models need a good set of features as an input, which are provided by image processing techniques in our case it is Contourlet Transform (CT) and Local Binary Pattern Histogram (LPBH). Deep learning is specialized form of ML that adopts neural networks to understand complex patterns in the data. Pre-trained convolution neural networks (CNNs) allow us to use state-of-the-art DL models by drastically reducing time and space complexity.

The success of any image classification task majorly relies in the efficient features' extraction and representation. In the proposed approach, we jointly employ machine learning and deep learning models for constructing the final voting classifier for the binary classification of the skin cancer detection. In the machine learning side, logistic regression and linear SVM are utilized for the classifying skin cancerous images. Input features for these two classifiers are synthesized from the state-of-the-arts image processing techniques such as contourlet transform and local binary pattern histogram followed by PCA for dimensionality reduction. In the deep learning, we experiment several pretrained CNNs along with some customization in the architectures for skin cancer classification. Finally, outputs of these two ML classifiers and best pretrained CNN are combined to apply voting classifier.

1.1 Contribution of research work

In this paper, we designed a novel approach for classification of benign and malignant melanoma. The main contributions of the paper are as follows:

- The paper introduces a combination novelty for efficient feature extraction. To our knowledge, combining automatic and manual feature extraction using image processing, machine learning and deep learning together has not been researched yet.
- Firstly, we tackle the problem of meaningful features' extraction for skin cancer classifier. We have utilized broadly two methods for feature extraction i.e. manual and automatic by using machine learning and deep learning respectively.
- In manual features' extraction, we adopted image processing techniques such as contourlet transform and LBP histogram, helped in recognising various features such as borders, contrast changes, and shapes. We experimented various ML models that map the extracted features to the class labels. We selected two top performing ML models i.e. Logistic Regression (LR) and Linear Support Vector Machine (LSVM).
- For automatic feature extraction, we employed DL models based on transfer learning technique. We trained state-of-the-art pretrained models such as VGG19, InceptionV3, ResNet, and DenseNet, after the process of image pre-processing and data augmentation. Among these pretrained CNNs, we select the top performer i.e. VGG19 to make the ensemble of ML/DL model. Moreover, recent works [4], have demonstrated the efficiency of using CNNs for feature extraction and SVMs for classification for images, which lead to our choice of these combination of models. SVM perform well for the classification tasks when provided efficient features as an input which is extracted using the fine-grained features' extractor such as VGG.
- To do a final classification, we combined the predictions of DL and ML models by using a voting approach. In this regard, a majority voting approach is applied. Here, total 3 classifiers are utilized for classification of skin cancer, discussed in Section 3. The prediction which is supported by more than 2 out of 3 classifiers is considered as the final classification.
- The approach is successfully implemented and benchmarked against a publicly available skin lesion dermoscopic image dataset (ISIC Archive). Proposed model outperforms the existing ML/DL models with higher accuracy.

The remaining paper is structured as follows; Section 2 provides the existing related works in the field of skin cancer classification using various ML/DL models. We analysed various methods applied on the skin cancer classification problem, primarily based on manual and automatic features' extraction techniques. Section 3 presents the proposed framework, and discusses the methodology applied by the ML and DL models. Section 4 presents the experimental results of the ML, DL, and ensemble model. In addition, this section also presents the comparative study between the proposed model and the contemporary state-of-the-arts models for skin cancer detection. Lastly, Section 5 concludes the paper and presents future research directions.

2 Related works

The recent advances in the field of computer vision and deep learning have achieved significant height in the field of prediction. Various computer vision tasks such as cancer

detection, self-driving cars, segmentation of medical images, and many more are getting leveraged from the huge improvements in the field of deep learning and related subfields [11, 12]. Skin cancer detection is not an exception of the same but suffering from the availability of large size well annotated dataset for the underlying task. Current approaches for the skin cancer classification involve two lines of approaches such as handcrafted features and DL for automated features extraction. Using handcrafted features, image processing techniques are applied to extract relevant features which are later provided to a ML model to learn from these features.

Recent works using handcrafted features extraction is reviewed in this section. Elgamal and Mahmoud [15] utilized feature extraction techniques namely wavelet transformation. The features extracted are further subjected to dimensionality reduction before being used for classification task. The k-nearest neighbour (k-NN) and artificial neural network (ANN) are applied for classification of skin cancer that are based on clinical findings and correlate certain characteristics present in dermoscopic images and tumor depth. Around 81 descriptions based on parameters such as color, texture, shape, and pigment network features are extracted. A combination of logistic regression and neural networks is adopted to classify, with an overall accuracy of 95%.

In [37], a grey-scale morphology is utilized to avoid distortion of information due to unwanted obstructions like hair particles in the images. Three different algorithms are developed to segment a lesion namely, global thresholding, dynamic thresholding, and a 3-D color clustering concept. Furthermore, in order to calculate simple color features, original images are transformed from RGB scale to HSI space. Different supervised learning algorithms are implemented for classification of skin cancer, with the highest overall accuracy of 77.6%. In [17], melanoma skin cancer classification is proposed using support vector machines (SVM) as the classifier. Before feature extraction, an image pre-processing and segmentation using thresholding is performed. Here, feature extraction algorithm utilized Gray Level Co-occurrence Matrix (GLCM) and four dermoscopic features are given importance represented by the acronym ABCD (Asymmetry, Border, Colour, Diameter). Further, dimensionality reduction is performed using Principle Component Analysis (PCA) technique, hence the proposed technique achieved an overall accuracy of 92.1%.

Subsequently, various supervised learning techniques are studied and analysed in [1]. SVM, ANN and AdaBoost algorithms are compared on the task of classification of benign and malignant melanoma images. Primarily, image pre-processing techniques are applied such as de-noising, sharpening of image, reshaping, and segmentation for further processing. Segmentation is accomplished using k-means clustering algorithm, here SVM and AdaBoost algorithms achieved the best performance on the dataset under experimentation.

Eventually, DL for automated feature extraction is investigated wherein neural network is employed to automatically filter out important features and learn complex relationships in images for classification. In [21], an automatic skin lesion classification technique is proposed, Ph2 dataset is utilized for training, consisting of 200 skin cancer images. To curb the problem of a small dataset, data augmentation is applied to increase the image size to 6600 by rotating the images. Transfer learning technique based on pretrained architecture of AlexNet is employed wherein weights are updated for the neural network and stochastic gradient descent (SGD) algorithm is applied. Moreover, the performance of the model is evaluated using the metrics such as accuracy, precision, sensitivity and specificity. The model is shown to outperform other existing methods.

Zhang et al. [54] discussed the shortcomings of deep convolution neural networks (DCNN) for the task of classifying skin lesions. The inability of DCNN to focus on semantically

meaningful parts is highlighted. Hence, attention residual learning CNN is proposed. The network is composed of multiple attention residual learning blocks, which uses residual learning to discriminate between the images. The network is tested on ISIC-skin 2017 dataset and outperforms other state-of-the-art methods. In [50], an ensemble of neural networks is developed for melanoma classification, for classification dermoscopic datasets are utilized i.e. xanthous race dataset and Caucasian race dataset. Colour, border, and texture features are broadly considered for classification. Finally, the lesion objects are classified using a neural network ensemble moreover to improve the performance, Back propagation neural network (BPNN) is combined with fuzzy neural networks (FNN).

In [52], deep convolution neural networks are employed to achieve automatic melanoma recognition. Moreover, Residual learning is deployed to prevent overfitting, which is often associated with deep neural networks. To achieve accurate segmentation, multi-scale contextual information integration scheme is adopted. A two-stage network is devised by combining the residual network and deep convolution network. The proposed framework is evaluated on the ISBI 2016 Skin Lesion Analysis Towards Melanoma Detection Challenge, and an area under curve (AUC) score of 80.5% was achieved.

In [24], numerous pre-trained neural network architectures were trained on the HAM10000 dataset, which consisted of seven different types of skin lesions. The images were resized to 224 × 224 and subjected to a due process of data augmentation. The authors tested several pretrained deep learning models, out of which ResNet50 model is shown to give the highest accuracy. For boosting the accuracy, an ensemble approach is implemented in the paper. By combining ResNet50, VGG16 and DenseNet model, an accuracy of 84% was achieved. Subsequently, Chaturvedi et al. [7], presented the DCNN for multi-class skin cancer detection on seven classes of skin lesions consisting in HAM10000 dataset. The best accuracy achieved for individual model i.e. ResNeXt101 is 93.20% whereas higher accuracy of 92.83% is reported for ensemble model as compared to state-of-the-art models.

In [50], the authors designed a three-step process for classifying cancer tumor as benign or malignant. Initially, a self-generating network is employed which is responsible for extracting lesions afterwards descriptive features are extracted. The lesion objects are finally classified using a neural network ensemble model, the model is benchmarked on the Caucasian race dataset and Xanthous race dataset. Moreover, Back propagation neural network and fuzzy neural networks were combined to achieve high performance.

Deep learning is characterized for automatic features extractor and followed by classification in case of skin cancer detection. However, several machine learning algorithms outperform the deep models with the help of several image pre-processing techniques. Traditional machine learning algorithms such as logistic regression and SVM also produces the tremendous results if the well-engineered features are provided as the input. Moreover, the proposed model is lighter in comparison with deep models. This is the main motivation behind the proposed method and it can be proven using the obtained experimental results.

The summarization of aforementioned literature is presented in the Table 1.

3 Materials and methods

In this section, we present underlying dataset and characteristics, features' extraction using image processing algorithms for providing better features as input to the machine learning models, pretrained CNNs for transfer learning and automatic features' extraction for the deep

Table 1 Summary of various skin cancer detection works

Ref.	Model Used	Contributions	Dataset	Results
[15]	k-NN and ANN	Wavelet transformed applied for feature extraction Dimensionality reduction to achieve good features	40 images	Accuracy −95%
[37]	global thresholding dynamic thresholding 3-D color clustering	feature extraction by clinical findings, features are correlated directly by melanoma thickness	250 dermoscopic images [3]	Accuracy – 77.6%
[17]	k-NN	binary mask identification for skin lesion using various segmentation techniques	5393 images	Sensitivity - 87% Specificity - 92%
[1]	SVM, ANN and AdaBoost	all features are fused for final classification Thresholding is applied for segmentation, statistical feature extraction through GLCM, PCA employed for feature selection, classification is performed by SVM	ISBI 2016	Accuracy – 92.1%
[21]	AlexNet	pretrained deep model and transfer learning augmentation and fine tuning the parameters	Ph2 dataset	Accuracy – 98.61%
[54]	Residual CNN	multiple attention residual blocks	ISIC-Skin 2017	Accuracy – 85%
[50]	Ensemble of neural network i.e. BPNN and FNN	Colour, border, and texture featured extracted using neural network	Xanthous race caucasian race	Accuracy - 94.17%
[52]	DCNN	Residual learning multi-scale contextual information integration	ISBI 2016	AUC 80.5%
[24]	ResNet50, VGG16 and DenseNet	Pretrained model for feature extraction Model ensembling for better feature extraction	HAM10000	Accuracy – 80%
[7]	DCNN	Multi-class skin cancer detection Model ensembling	HAM10000	Accuracy - 93.20%

Table 2 Distribution of image in dataset

Image Class	Training Set	Test Set
Benign	1440	401
Malignant	1197	259

learning models, and final voting classifier for the skin cancer classification. Moreover, popular images processing techniques i.e. contourlet transform and local binary pattern histogram are also explained in brief.

3.1 Dataset

For the underlying experimentations, we utilized Kaggle skin cancer dataset [26] consisting of processed skin cancer images of ISIC Archive dataset [13]. The dataset consists of total 2637 training images and 660 testing images, each of these having a resolution of 224 × 224. It is a binary skin cancer dataset consisting of two main classes viz. *malignant* and *benign*. The Table 2 shows the distribution of images into benign and malignant class in the dataset.

3.2 Machine learning approach

Machine learning is a field of computer science that allows systems to learn from data without being explicitly programmed. The various ML algorithms such as SVM and logistic regression are quite popular for the classification and regression tasks. Binary SVM is based on separating the datapoints in two classes with the help of a separating hyperplane, generating maximum geometric and functional margin. However, logistic regression is another important linear ML classifier utilizing the sigmoid function for the binary classification task with the help of a predefined threshold. There is always a trade-off in applying the linear kernel SVM or logistic regression, depends upon the nature and characteristics of the dataset. Herein, we employ both i.e., linear kernel SVM and logistic regression for generating votes in the final skin cancer classification. Before utilizing these ML classifiers for the skin cancer detection, we employ image processing algorithms viz. contourlet transform and local binary pattern histogram for better features' representation. Output features from these two image processing algorithms are provided as inputs to the ML classifier after performing dimensionality reduction using PCA. We explain these image processing algorithms and PCA in brief in the section.

DL methods do not allow us to manually extract features from the images whereas the feature extraction process is completely automated [30]. Hence, in the cases where handcrafted features might increase the accuracy, deep learning suffer from a setback. ML models allow to provide features after performing extensive features' extraction.

3.2.1 Contourlet transform

Contourlet Transform is one of the image transformation techniques utilized widely for feature extraction from the visual input. The important and dominant features in an image are easily extracted by CT. The discrete contourlet transform has a fast iterated filter bank algorithm [14] that requires an order N operation for N-pixel images. CT is highly preferred for generating features from images because of its possession of an extensive set of properties such as multi-resolution, localization, critical sampling, directionality, and anisotropy. If the original signal
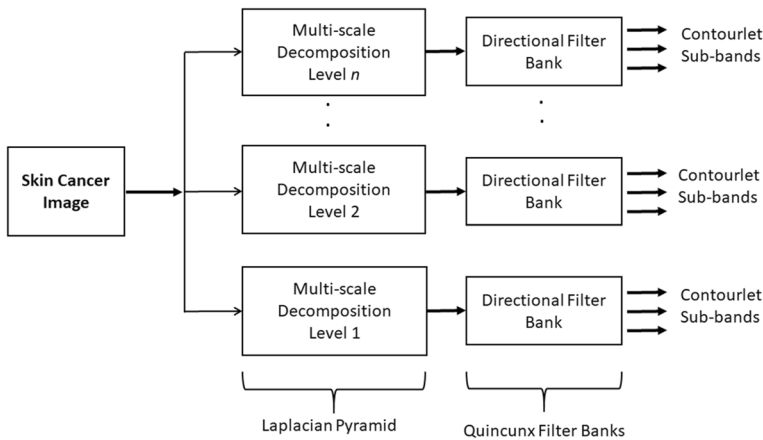
Fig. 2 Image Decomposition and features' extraction using Contourlet Transform

or image is piecewise smooth, wavelet transform is natural choice for edge details acquisition. In the context of skin cancer detection, the edges are not always piecewise smooth and situated across fine contours. Hence, CT is preferred for edge contour information acquisition [35]. Figure 2 shows the decomposition of image wherein CT utilizes Laplacian pyramid for multi-scale decomposition and on each sub-band, directional filters are employed.

3.2.2 Local binary pattern histogram

Ojala et al. [34] proposed the local binary patterns technique for efficient texture classification. Eventually, LBP have become increasingly popular for computer vision applications as well. The operations of LBP are computationally simple and efficient in describing local structures in images [22]. LBP for a pixel is represented by a decimal number for processing the neighbourhood of the pixel, can be expressed using Eq. 1. Formally, given a pixel (x_c, y_c), the resulting LBP can be computed in decimal form as follows [13]:

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} S(i_p - i_c) \times 2^p \tag{1}$$

where i_c and i_p are the gray-level values respectively of the central pixel and P surrounding pixels in the circle neighbourhood with a radius R, and function $s(x)$ is defined using Eq. 2 [26]:

$$s(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} \tag{2}$$

3.2.3 Dimensionality reduction

Features obtained after applying contourlet transform and LBP histogram are concatenated for providing inputs to the ML classifiers. Out of these extracted features, the most important and dominating features are selected by applying the PCA method. The rest of the features are just ignored in the classification process.

Principal component analysis PCA is a dimensionality-reduction method that is often applied to reduce the dimensionalities of large datasets, by transforming a large set of features into a smaller set that still represents the features effectively. Mathematically, PCA depends on – 1) eigen-decomposition of positive semi-definite matrices and 2) singular value decomposition (SVD) of rectangular matrices [33]. It is determined by eigen-vectors and eigen-values. Using PCA, we convert the initial image of 224 × 224 into a set of 50 data points. These 50 data points are extracted from the concatenation of contourlet transform output and LBP histogram output.

3.2.4 Training ML models

Once the contourlet transform and LBP extracts the most meaningful features from the input images, these features are provided as an input to the ML models to learn the features after performing dimensionality reduction using PCA. Figure 3 presents the flow of ML model adopted after applying CT and LBP for features extraction, finally LR and LSVM are employed to classify the input image. In this work, initially, we trained five ML models on the ISIC dataset and choses the top two performing models i.e. LR and LSVM. Performance of ML models on the test set of 660 images are displayed in results section.

3.3 Deep learning approach

To make our model robust, in this work, we combine the ML and DL models to classify skin cancerous images. This allows us to combine the benefits of handcrafted features' extraction and automatic features' extraction using ML and DL models respectively. While preparing the
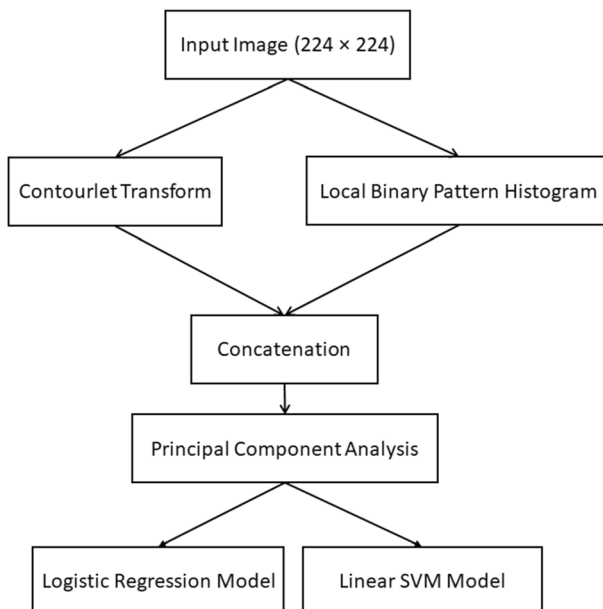

Fig. 3 Flow of ML Model

deep learning model for skin cancer detection, dataset is first pre-processed and augmented for better model performance. We utilized five pretrained CNNs viz. InceptionV3, VGG19, ResNet50, DenseNet201, and InceptionResnetV2 for transferring the learning in skin cancer classification. We briefly describe data pre-processing and augmentation, transfer learning, and these pretrained CNNs in this section.

3.3.1 Data augmentation

One of the common problems in ML/DL models is model overfitting [51], especially in case of smaller datasets. Overfitting means that a model performs well on the given train dataset but fails to achieve the similar performance on slightly different images (i.e. test dataset). Thus, to increase the dataset for model training, we adopt the various data augmentation [40] techniques such as rotation, horizontal and vertical flipping, and brightness and altering contrast. Image augmentation is performed using the ImageDataGenerator function provided in the Keras library.

In the training set, the total number of images are 2637 images, with 1440 images are in benign class and 1197 images are in malignant class. To overcome the challenges of relatively
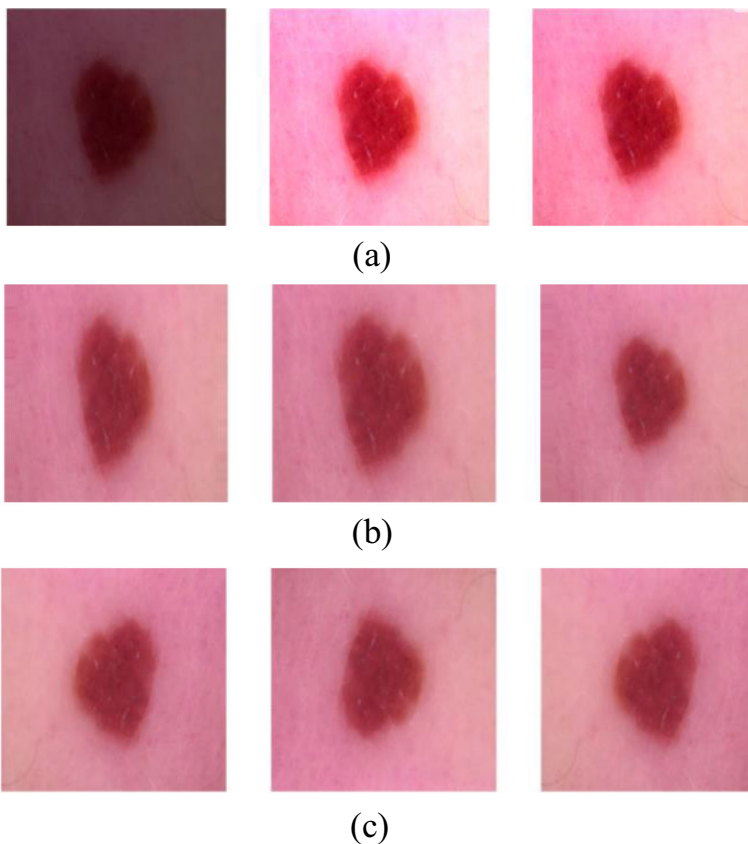

Fig. 4 a Augmentation based on varying brightness, **b** Augmentation based on zooming on image, and **c** Augmentation based on horizontal and vertical flipping of image

small dataset we used image augmentation, which supplies the model unique variants in each epoch. The model is trained for 50 epochs, and the training data is split into 70% training and 30% validation data. ImageDataGenerator in Keras provides the model augmented training images in each epoch. Test set comprises of 660 images with 401 benign images and 259 malignant images, respectively. The batch size for training is selected as 64. We have trained the various models on the dataset to measure the performance for skin cancer detection. Figure 4 presents the sample results of augmentation techniques.

3.3.2 Convolution neural networks

Convolution neural networks [18] are a form of deep neural networks utilized for automatic feature extraction. We adopted CNN over general artificial neural network because of the superior performance of CNN on image feature extraction [39]. CNNs deploy the use of mathematical operations called convolution over matrices of pixels. Combined with the use of filters, which help in extracting spatial meaning of various components in an image, CNN tend to outperform other types of deep neural networks on image data.

3.3.3 Transfer learning

Transfer learning [44] offers us to reutilize the learned parameters by a neural network, moreover, it felicitates to apply highest performing models on the most popular image classification dataset i.e. '*ImageNet*', containing millions of images and can be modified as per domain and problem requirements. We directly employed the pretrained model /problem after performing changes to the input size, replacing the number of output classes (in our case it is 2) as per requirement instead of 1000 classes i.e. in the original ImageNet problem. Moreover, we have added some additional layers to customize the model to best fit for skin cancer detection problem. Instead of directly employing pretrained model, another option available is training the entire deep learning model from scratch. However, the deep learning model trained with ImageNet weights gave a better performance on the dataset as compared to the model from scratch, which is shown in the results section. We experimented five powerful deep neural network architectures for this work and chosen the best performing model to construct ensemble with our proposed ML approach.

The input images are resized as 224 × 224, the learning rate of 0.0001 is utilized to train the CNN models, Adam as an optimizer is applied, and some customization is suggested for the models under consideration for skin cancer detection.

3.3.4 InceptionV3

InceptionV3 [42] is a convolutional neural network architecture from the Inception family that makes several improvements such as Label Smoothing, smart Factorized convolutions, and the use of an auxiliary classifier to propagate label information lower down the network. Google's InceptionV3 architecture is re-trained on our dataset by fine-tuning across all layers and replacing top layers with one average pooling, four fully-connected, and finally the sigmoid activation function in the output layer allowing to classify 2-diagnostic categories i.e. *benign* or *malignant*. The input images are resized to 224 × 224 hence the images become compatible with this pretrained model.

3.3.5 VGG19

As the name suggests, VGG19 is a pretrained CNN model consists of 19 layers. The trained VGG19 [41] architecture is originally trained on the ImageNet dataset with 1000 classes. In this work, VGG19 architecture is experimented and re-trained on our dataset by fine-tuning across all the layers and replacing top layers with four fully-connected layers, and lastly the sigmoid layer is applied to achieve the classification result. Input is made compatible by resizing it as per the original input size of the VGG19.

3.3.6 ResNet

ResNet50 [19] architecture is re-trained on the dataset for experimentation and fine-tuning across all the layers are applied. The top layers are replaced with four fully-connected layers with 'relu' as the activation function, and sigmoid layer is applied at last to comprehend the results in 2 diagnostic categories. Here, the network uses identity mapping for resolving the vanishing gradient problem and to train much deeper networks. This identity mapping does not have any parameters but it adds the output from the previous layer to the layer ahead. The Skip Connections between layers adds the outputs from previous layers to the outputs of stacked layers [10]. The ResNet50 architecture is kept as it is with their pretrained weights. Similar to the previous pretrained CNNs, four fully connected layers have been added at the top with the activation function as 'relu'. Sigmoid function is used as the activation function for the final output layer. For the purpose of extensive testing, we also implemented another variant of ResNet, namely ResNet152v2 which contains a similar architecture to ResNet50, with 152 layers in the model.

3.3.7 DenseNet201

DenseNet201 [42] is a convolution neural network which is 201 layers deep network. Each layer in DenseNet201 receives knowledge from the previous layers. To solve the vanishing gradient problem, this architecture uses a simple connectivity pattern to ensure the maximum flow of information between layers both in forward and backward computation. The layers are connected in a way such that the inputs from all the preceding layers passes through its own feature-maps to all the subsequent layers. To facilitate the down-sampling in the architecture, the entire architecture is divided into multiple densely connected blocks. The layers between these dense blocks are transition layers which perform convolution and pooling operations. On top of this pre-trained architecture, four fully connected layers are added with 'relu' as activation function. Classification is done at the final output layer using sigmoid as the activation function.

3.3.8 InceptionResnetV2

This model is the combination of Inception and ResNet [23] models. InceptionResNetV2 architecture is re-trained on our dataset and fine-tuning is proposed across all the layers - top layers replaced with four fully-connected layers, and sigmoid layer is applied at the end to classify the image in 2-diagnostic categories.

Hence, all the above models are trained for the binary classification of skin cancer images on ISIC dataset, and it is found that the VGG19 gives comparatively higher performance
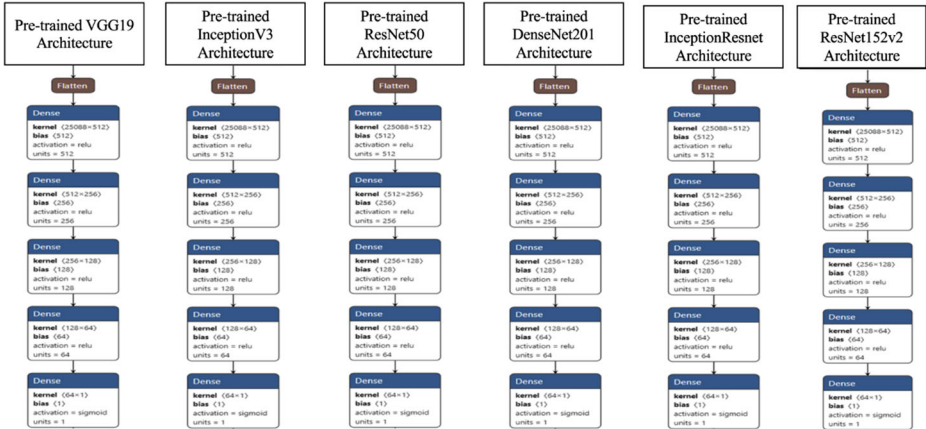
Fig. 5 Pretrained CNN architectures for skin cancer detection after being customized with dense and sigmoid layer

amongst all the pretrained CNNs. The schematic of these pretrained CNNs along with the customization for the skin cancer classification are illustrated using Fig. 5.

3.3.9 Fine tuning of classification model

The loss function used is binary cross entropy. The equation for binary cross entropy is given as follows:

$$Loss = -\frac{1}{N} \sum_{i=1}^{N} y_i \cdot \log(p(y_i)) + (1-y_i) \cdot \log(1-p(y_i)) \tag{3}$$

Where y_i is the output label (1 for malignant class and 0 for benign) and $p(y_i)$ is the predicted probability of the output class. For faster minimization of the loss, Adam [28] optimizer is applied. It is based on adaptive estimation of first order and second order moments, that allows weight optimization in the most efficient way. Weights are updated according to Eq. 4.

$$w_t = w_{t-1} - \frac{\eta \cdot \widehat{m}_t}{\sqrt{\widehat{v}_t + \varepsilon}} \tag{4}$$

Where w_t is the weight at a given time step and η is the learning rate which is set 0.0001. $\widehat{m}_t$ and $\widehat{v}_t$ are given by Eqs. 5 and 6, also known as first and second moment estimates for efficient loss convergence.

$$\widehat{m}_t = \frac{m_t}{1-\beta_1^t} \tag{5}$$

$$\widehat{v}_t = \frac{v_t}{1-\beta_2^t} \tag{6}$$

Where β_1 and β_2, called as decay rates are set to 0.9 and 0.999 respectively.

The learning rate, optimizer and loss functions are kept same for all deep learning models tested.

3.4 Proposed ensemble model for skin Cancer detection

In this proposed approach, we construct the ensemble of ML and DL models using voting mechanism. The two ML models i.e. logistic regression and linear SVM are combined with the top performing pretrained CNN i.e. VGG19 for the final classification of skin cancer, schematic of the proposed approach can be illustrated using Fig. 6. The modified VGG19 is shown in Fig. 7 wherein the additional layers are appended to the original pretrained architecture of VGG19. Figure 7 also demonstrate the significant reduction in trainable parameters using pretrained VGG19.

Many widely utilized approaches in the field of image feature extraction problems indicate the benefits of using an ensemble of 2 deep learning models for feature extraction [9, 48, 49]. Thus, we also experimented using 2 DL models for the voting approach, namely VGG19 and Resnet152v2. However, the single DL model shown the higher performance metrics with a lower complexity as indicated in Table 7. Therefore, we proceeded with the approach of using 2 machine learning models along with VGG19 model.

By combining the deep learning and machine learning counterpart/models, we are able to increase the level of meaningful features' extraction and final classification. Since features' extraction is the main/major challenge in any classification, thus, our ensemble model makes use of two separate approaches for extracting features. While the deep learning model trains directly on images after pre-processing and augmentation, the ML models utilize the output from contourlet transform and LBPH to make for better features' representation.
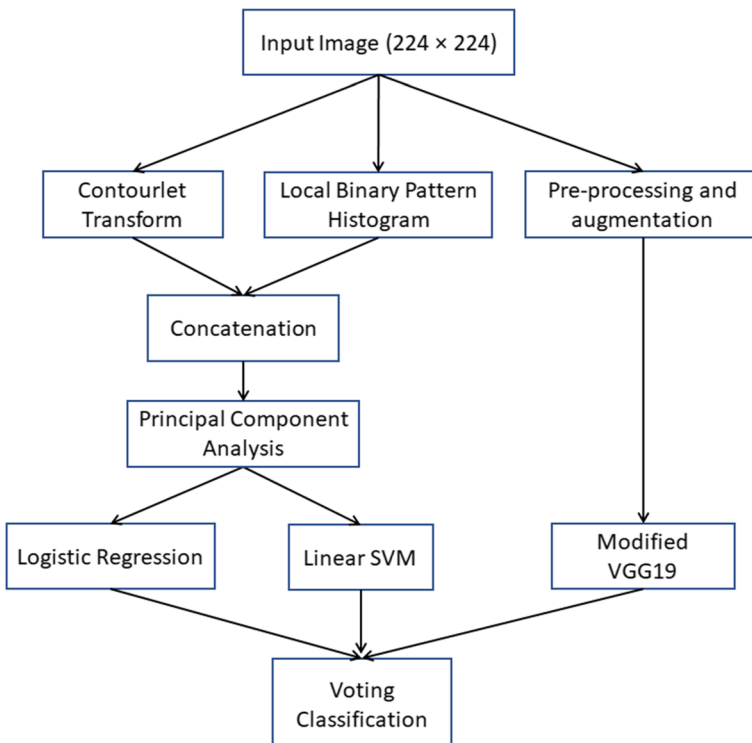

Fig. 6 Ensemble Model Flow in skin cancer detection

```
-------------------------------------------------------------------------
Layer (type)                  Output Shape              Param #
=========================================================================
vgg19 (Functional)            (None, 7, 7, 512)         20024384
-------------------------------------------------------------------------
flatten_1 (Flatten)           (None, 25088)             0
-------------------------------------------------------------------------
dense_5 (Dense)               (None, 512)               12845568
-------------------------------------------------------------------------
dense_6 (Dense)               (None, 256)               131328
-------------------------------------------------------------------------
dense_7 (Dense)               (None, 128)               32896
-------------------------------------------------------------------------
dense_8 (Dense)               (None, 64)                8256
-------------------------------------------------------------------------
dense_9 (Dense)               (None, 1)                 65
=========================================================================
Total params: 33,042,497
Trainable params: 13,018,113
Non-trainable params: 20,024,384
-------------------------------------------------------------------------
```

Fig. 7 Modified VGG19 architecture demonstrating reduction in the trainable parameters

The machine learning model on the other hand has made use of hand-crafted features provided applying a thorough process of image processing. The machine learning model then learns to map the relationship between the supplied features and the correct output. In our case, Logistic Regression and Linear SVM found to be the best performing ML models on this dataset, and these ML models are utilized for the experimentation.

By using voting approach, we are able to combine the effects of different feature extraction techniques. Voting allows to take into consideration the results of both DL and ML models, which represent automatic and manual feature extraction respectively. By this combination, we are able to achieve excellent performance of our ensemble model.

4 Results and discussions

To evaluate the proposed model, the following metrics are considered for the performance measurement.

Accuracy Precision quantifies the correct predictions out of the total number of samples.

$$Accuracy = \frac{True\ Positive + True\ Negative}{Total\ Number\ of\ Samples} \tag{7}$$

Precision Precision quantifies the number of positive class predictions that actually belong to the positive class.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \tag{8}$$

Recall: Recall quantifies the number of positive class predictions out of all the actual positive examples in the dataset.

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{9}$$

*F*1-*Score*: F1-Score measure provides a single score that provides the combined measure of the previous two measures i.e., precision and recall and can be defined as a harmonic mean of both of these.

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{10}$$

Support Support refers to the number of images used for each class in order to calculate evaluation metrics.

For the experimentations, we utilized Kaggle platform consisting of Linux operating system supported by Nvidia K80 GPU with Tensor Flow 2.4, Keras 2.4, and Python 3.7. We have chosen Kaggle platform due to the availability of extensive deep learning and machine learning library facilities and support for faster training of complex deep learning or ensemble models on the GPU.

We observed that the individually VGG19 model achieved an accuracy of 84% on the training dataset and an accuracy of 91% on the validation dataset. The model is trained for 50 epochs, Adam optimizer is applied, and binary cross entropy is used as the loss function. Moreover, other individual deep learning models such as InceptionV3, ResNet50, DenseNet201, InceptionResNetV2 are also experimented and tested for skin cancer detection. Table 3 shows the comparison of all these models wherein F1-score, precision, recall, and accuracy are measured. It is found that the VGG19 outperforms other individual models. Herein, we consider all the performance metrics for evaluation of the proposed model. Though, some of the models perform well on some of the aspects but could not draw consistent results across all the performance matrices. ResNet i.e., ResNet152v2 variant can be consider as the second best performer after VGG19 due to resolving the vanishing gradients problem of the deep model with the help of skip and residual connections. To perform hyper-parameter tuning, we experimented various values for parameters such as batch size, validation split, and varying optimizers along with the Keras Tuner. The model loss and accuracy are graphically presented in Fig. 8 for the most promising hyperparameter combination employed for the proposed model. The Keras Tuner library is utilized in our model to find the most optimal set of parameters for the learning algorithm. Moreover, we also did several experimentations with slightly varying hypermeters to validate the findings of Keras Tuner.

Table 3 Comparative analysis of pretrained deep learning models

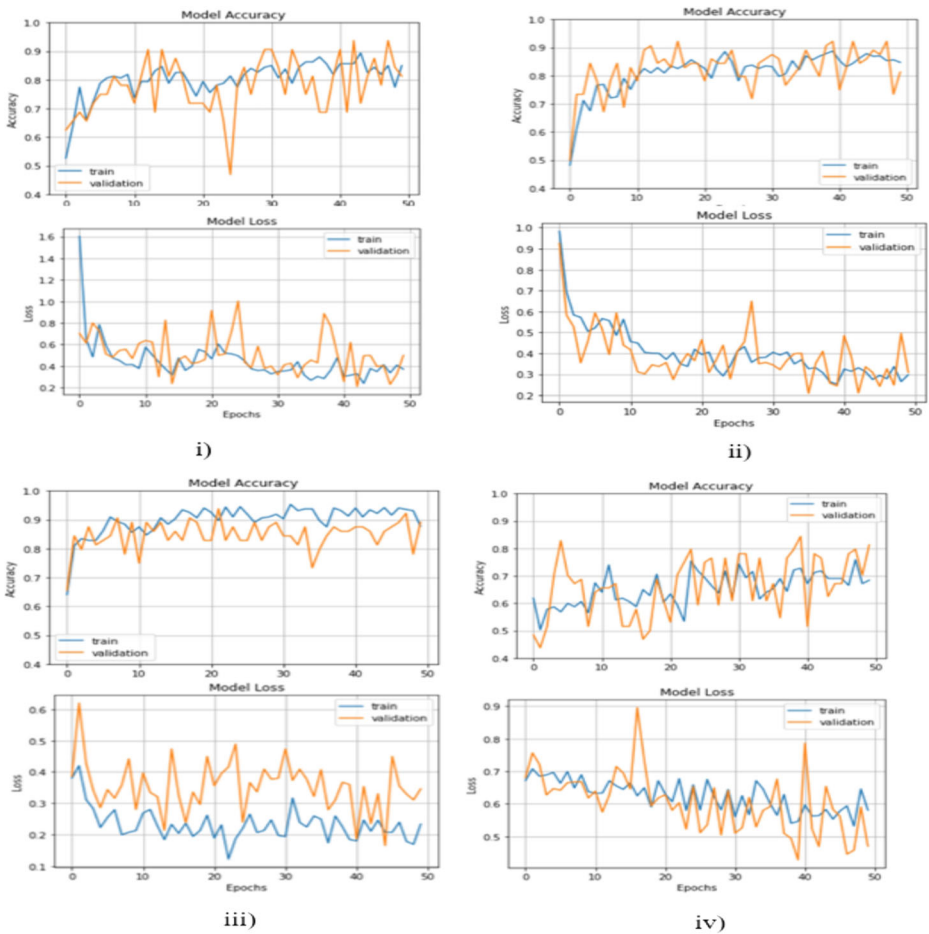| Model	Precision		Recall		F-1 Score		Accuracy
	Benign	Malignant	Benign	Malignant	Benign	Malignant	Overall
InceptionV3	0.61	0.89	0.87	0.66	0.72	0.76	0.74
ResNet50	0.50	0.94	0.91	0.61	0.65	0.74	0.70
DenseNet201	0.62	0.95	0.93	0.68	0.75	0.79	0.77
InceptionResnetV2	0.68	0.80	0.81	0.68	0.74	0.73	0.73
ResNet152v2	0.85	0.79	0.83	0.82	0.84	0.81	0.83
VGG19	0.82	0.86	0.87	0.80	0.85	0.83	0.84

Fig. 8 Modified VGG19 model's Accuracy vs Loss after training for 50 epochs with varying hyperparameters: i) Batch size = 32, optimizer = Adam, validation split = 0.1, ii) Batch size = 64, optimizer = Adam, validation split = 0.1, iii) Batch size = 64, optimizer = Adam, validation split = 0.3, and iv) Batch size = 64, optimizer = SGD, validation split = 0.3

As shown in Fig. 7, to train the VGG19 model it requires to handle large numbers of trainable parameter i.e., 33 million. Thus, by using pretrained weights, we have reduced the trainable parameters to 10.3 million, improving the time and space complexity drastically.

Using pretrained weights for VGG19 instead of training the complete model, lead to an increase in accuracy which is shown in Table 4. Transfer learning using pretrained CNNs can

Table 4 Comparative analysis of pretrained VGG19 (using ImageNet weights) and VGG19 (trained complete model) on test data

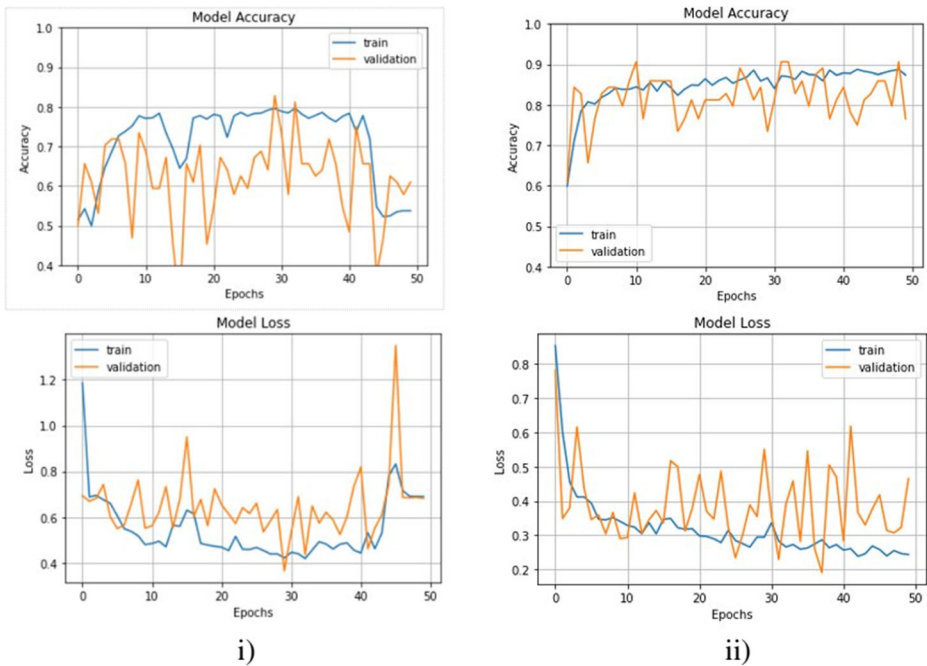| Model	Average Precision	Average Recall	Average F-1 Score	Accuracy
Modified VGG19 (trained complete model)	1.00	0.55	0.71	0.55
Modified VGG19 (using ImageNet weights)	0.82	0.87	0.85	0.84

Fig. 9 Modified VGG19 model's Accuracy vs Loss after training for 50 epochs i) Training Complete model, ii) Model with pretrained ImageNet weights and training only the fully connected layers at the end

be considered as best technique for better generalization in most of the computer vision tasks. Learning from scratch may lead to overfit the training data and suffer from bad generalization. From Table 4, training the complete model leads to an overfitting on one class, which is a result of using 33 million parameters of VGG19 (see Fig. 7) on a small sample of 2637 images in the training set. Figure 9 shows the graphical representation of accuracy and loss with epoch for the two models. Hence, for our framework, we utilize the modified VGG19 with ImageNet weights to achieve optimum performance, and trained the fully connected layers.

The ML models utilized for classification are also compared with other ML models. The testing performance metrics are presented in Table 5 wherein we observe that the Logistic regression is outperforming other machine learning models across all the performance matrices. Linear SVM is the second-best performer among all the machine learning models. We

Table 5 Testing Metrics Comparison for ML Models

Model	Precision		Recall		F1-Score		Accuracy
	Benign	Malignant	Benign	Malignant	Benign	Malignant	Overall
KNN	0.53	0.67	0.66	0.54	0.59	0.60	0.59
Decision Tree	0.73	0.68	0.73	0.68	0.73	0.68	0.71
SVM	0.84	0.79	0.83	0.80	0.83	0.79	0.82
Logistic Regression (Features' extracted using CT and LBPH)	0.99	0.78	0.85	0.99	0.91	0.88	0.90
Linear SVM (Features' extracted using CT and LBPH)	0.96	0.75	0.82	0.94	0.88	0.83	0.86

Table 6 Ensemble Model's Performance Metrics (2 ML Models and VGG19 + ResNet152v2)

	Precision	Recall	F1-Score	Support
Benign	0.99	0.85	0.91	421
Malignant	0.79	0.99	0.88	239
Overall F1-Score	–	–	0.90	660
Macro average	0.89	0.92	0.89	660
Weighted average	0.92	0.90	0.90	660

conclude that linear machine learning algorithms are performing better for skin cancer binary classification problem in comparison with nonlinear machine learning algorithms such as decision trees. Non-linear ML algorithms perform better on training data but may suffer from overfitting problem, however, this will not be the case with linear ML algorithms. Moreover, the evaluation metrics of the ensemble model with 2 ML models and 2 DL models is shown in Table 6. The ensemble model with 2 ML models and 1 DL Model (Modified VGG19) outperformed the other approaches is presented in Table 7. As we achieved improved performance metrics in this binary classification task on ISIC skin cancer dataset, the proposed model shall surely perform well on other binary skin cancer datasets as the underlying dataset utilized in this research work is balanced, robust and well annotated.

The proposed model offers overall F1-Score of 93% with an impressive individual recall score of 89% and 99% for the benign and malignant class, respectively (see Table 7). Moreover, the test set accuracy was found to be 93%. Further, Table 8 presents the comparative analysis of our ensemble model with other state-of-the-art methods for skin cancer classification. Hence, we can depict from the Table 8 that the proposed model outperforms the existing state-of-the-art ML/DL models for the skin cancer detection. Our model not only gives the highest test accuracy but also gives significantly higher results for other evaluation metrics like precision, recall and F1-score as well. Moreover, Table 8 shows the overall test accuracy has been increasing from 2016, it is due to the better deep learning models and emergent image processing techniques. Our model outperforms other approaches by a significant margin, not because of a more complex learning model but because of a more efficient method of features' extraction and final voting classifier. By combining both automatic and manual features' extraction techniques in DL and ML based models respectively, our models have been able to learn complex relationships between input features and output labels.

Faster model inference and better model performance metrics are considered to be the main objectives of any prediction task using ML/DL models. However, increased model complexity leads to better model performance metrics and slower model inference time, producing a trade-off between model complexity and model inference time. Though, the model complexity is slightly increased by introducing the ensemble of image processing, ML, and DL for this

Table 7 Proposed Ensemble Model's Performance Metrics (2 ML Models and VGG19)

	Precision	Recall	F1-Score	Support
Benign	0.99	0.89	0.94	401
Malignant	0.86	0.99	0.92	259
Overall F1-Score	–	–	0.93	660
Macro average	0.93	0.94	0.93	660
Weighted average	0.94	0.93	0.93	660

Table 8　Comparison of proposed model with state-of-the-art models

Ref.	Accuracy	Precision	Recall	F1-Score
Kawahara et al. [27] (2016)	79.5%	–	–	–
Lopez et al. [32] (2017)	81.3%	0.7974	0.7866	–
Brinker et al. [6] (2019)	83.9%	0.709	0.56	–
Wu J et al. [47] (2020)	85.7%	0.756	0.816	–
Farooq et al. [16] (2020)	86%	0.8	0.89	0.84
Proposed Model	93%	0.92	0.94	0.93

binary skin cancer classification task, the slight increase in the model inference time is counter balanced by the considerable improved performance in the classification task.

5 Conclusions and future works

In this paper, we explore a novel approach of extracting features from the images for skin cancer detection. We employed the combination of ML technique such as contourlet transform and LBP histograms, and VGG19 a deep learning model to extract the features. The features from these models are combining to achieve the final skin cancer classification using voting mechanism. The primary reason our model performs well is because of a better design of feature extraction. It has been seen in image classification problems that as long as feature extraction is not efficient and meaningful, no number of complex models can fit to the data. Hence, for the purpose of our study we leveraged the power of deep learning models and allowed them to extract meaningful features. On the other hand, we combined manual feature extraction by using powerful image processing techniques. The ensemble of this approach outperforms various state-of-the-art techniques in terms of overall test accuracy obtained on the ISIC Archive dataset. Despite the high performance of our model, it is not meant to replace dermatologists and radiologists. Instead, our model can serve the purpose of assisting doctors, helping drastically reduce the number of dales negatives which is crucial in medical diagnosis. Hence, we highly recommend the model in assisting dermatologists to diagnose various forms of skin cancer.

Moreover, this approach can be extended for similar image classification problems for enhanced feature extraction. In [38], demonstrated how a binary mage classification algorithm can be applied to offer high performance on similar datasets. In this paper, we presented the use of ensemble learning for 2D images or data. The same concept can be extended for 3D multimedia applications [8, 31] as well. Various feature search optimization techniques [46, 53] can be applied to aid feature extraction in a 3D space. Furthermore, the proposed model can be utilized to assist dermatologists for better diagnosis, instead of replacing them. More specifically, we can integrate the proposed model on the smartphones and web applications to provide easy interface for doctors [20], without possessing advanced knowledge of the working of individual models in depth. The aforementioned points are the future research direction in the field of skin cancer detection.

Data availability　Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Declarations

Competing interests   The authors of this publication declare there is no conflict of interest.

References

1. Alquran H, Qasmieh IA, Alqudah AM, Alhammouri S, Alawneh E, Abughazaleh A, Hasayen F (2017) The melanoma skin cancer detection and classification using support vector machine. In IEEE Jordan conference on applied electrical engineering and computing technologies (AEECT), 1-5
2. American Cancer Society - Key Statistics for Melanoma Skin Cancer (2021) https://www.cancer.org/cancer/melanoma-skin-cancer/about/key-statistics.html. Accessed on 12th January 2021.
3. Argenziano G, Soyer HP, De Giorgio V, Piccolo D, Carli P, Delfino M, Wolf IH (2000) Interactive atlas of dermoscopy
4. Basly H, Ouarda W, Sayadi FE, Ouni B, Alimi AM (2020) CNN-SVM learning approach based human activity recognition. In: Proceedings of international conference on image and signal processing. Springer, Cham, pp 271–281
5. Brinker TJ, Hekler A, Enk AH, Berking C, Haferkamp S, Hauschild A, Weichenthal M, Klode J, Schadendorf D, Holland-Letz T, Kalle CV, Fröhling S, Schilling B, Utikal JS (2019) Deep neural networks are superior to dermatologists in melanoma image classification. Eur J Cancer 119:11–17
6. Brinker TJ, Hekler A, Enk AH, von Kalle C (2019) Enhanced classifier training to improve precision of a convolutional neural network to identify images of skin lesions. PLoS One 14(6):e0218713
7. Chaturvedi SS, Tembhurne JV, Diwan T (2020) A multi-class skin Cancer classification using deep convolutional neural networks. Multimed Tools Appl 79(39):28477–28498
8. Chen Y, He F, Wu Y, Hou N (2017) A local start search algorithm to compute exact Hausdorff distance for arbitrary point sets. Pattern Recogn 67:139–148
9. Chen C, Wang G, Peng C, Fang Y, Zhang D, Qin H (2021) Exploring rich and efficient spatial temporal interactions for real-time video salient object detection. IEEE Trans Image Process 30:3995–4007
10. Chollet F (2017) Xception: deep learning with depthwise separable convolutions. In proceedings of the IEEE conference on computer vision and pattern recognition, 1251-1258
11. Chowdhary CL, Alazab M, Chaudhary A, Hakak S, Gadekallu TR (2021) Computer vision and recognition systems using machine and deep learning approaches: fundamentals, Technologies and Applications. Inst Eng Technol
12. Chowdhary CL, Reddy GT, Parameshachari BD (2022) Computer vision and recognition systems: research innovations and trends. CRC Press
13. Codella NC, Gutman D, Celebi ME, Helba B, Marchetti MA, Dusza SW, Halpern A (2018) Skin lesion analysis toward melanoma detection: a challenge at the 2017 international symposium on biomedical imaging (ISBI), hosted by the international skin imaging collaboration (ISIC). In IEEE 15th international symposium on biomedical imaging (ISBI 2018), 168-172
14. Do MN, Vetterli M (2005) The contourlet transform: an efficient directional multiresolution image representation. IEEE Trans Image Process 14(12):2091–2106
15. Elgamal M (2013) Automatic skin Cancer images classification. Int J Adv Comput Sci Appl 4(3):287–294
16. Farooq MA, Khatoon A, Varkarakis V, Corcoran P (2020) Advanced deep learning methodologies for skin Cancer classification in prodromal stages. arXiv preprint arXiv:2003.06356
17. Ganster H, Pinz P, Rohrer R, Wildling E, Binder M, Kittler H (2001) Automated melanoma recognition. IEEE Trans Med Imaging 20(3):233–239
18. Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, Liu T, Wang X, Wang L, Wang G, Cai J, Chen, T (2015) Recent advances in convolutional neural networks. CoRR, arXiv preprint arXiv
19. He K, Zhang X, Ren S, Su, J (2016) Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, 770–778. https://doi.org/10.1109/CVPR.2016.90
20. Hebbar N, Patil HY, Agarwal K (2020) Web powered CT scan diagnosis for brain hemorrhage using deep learning. In 2020 IEEE 4th conference on Information & Communication Technology (CICT), 1–5
21. Hosny KM, Kassem MA, Foaud MM (2018) Skin cancer classification using deep learning and transfer learning. In 9th Cairo international biomedical engineering conference (CIBEC), 90-93. https://doi.org/10.1109/CIBEC.2018.8641762
22. Huang D, Shan C, Ardabilian M, Wang Y, Chen L (2011) Local binary patterns and its application to facial image analysis: a survey. IEEE Trans Syst Man Cybern Part C Appl Rev 41(6):765–781

23. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In proceedings of the IEEE conference on computer vision and pattern recognition, 4700-4708
24. Ismail MA, Hameed N, Clos J (2021) Deep learning-based algorithm for skin Cancer classification. In proceedings of the international conference on trends in computational and cognitive engineering, 709–719
25. Jana E, Subban R, Saraswathi S (2017) Research on skin Cancer cell detection using image processing. In IEEE international conference on computational intelligence and computing research (ICCIC), 1–8. https://doi.org/10.1109/iccic.2017.8524554
26. Kaggle Skin Cancer Dataset (2021) https://www.kaggle.com/fanconic/skin-cancer-malignant-vs-benign. Accessed on 7 February 2021.
27. Kawahara J, Hamarneh G (2016) Multi-resolution-tract CNN with hybrid pretrained and skin-lesion trained layers. In: International workshop on machine learning in medical imaging. Springer, Cham, pp 164–171
28. Kingma DP, Ba J (2014) Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980
29. Labani S, Asthana S, Rathore K, Sardana K (2020) Incidence of melanoma and nonmelanoma skin cancers in Indian and the global regions. 17:906–911. https://doi.org/10.4103/jcrt.JCRT_785_19
30. Liang H, Sun X, Sun Y, Gao Y (2017) Text feature extraction based on deep learning: a review. EURASIP J Wirel Commun Netw 2017(1):1–12
31. Liang Y, He F, Zeng X (2020) 3D mesh simplification with feature preservation based on whale optimization algorithm and differential evolution. Integr Comput-Aided Eng 27(4):417–435
32. Lopez AR, Giro-i-Nieto X, Burdick J, Marques O (2017) Skin lesion classification from dermoscopic images using deep learning techniques. In 13th IASTED international conference on biomedical engineering (BioMed), 49-54
33. Mishra SP, Sarkar U, Taraphder S, Datta S, Swain DP, Saikhom R, Laishram M (2017) Multivariate statistical data analysis-principal component analysis (PCA). Int J Livestock Res 7(5):60–78. https://doi.org/10.5455/ijlr.20170415115235
34. Ojala T, Pietikäinen M, Harwood DA (1996) Comparative study of texture measures with classification based on featured distributions. Pattern Recogn 29(1):51–59
35. Patil HY, Kothari AG, Bhurchandi KM (2016) Expression invariant face recognition using local binary patterns and contourlet transform. Optik 127(5):2670–2678
36. Riker AI, Zea N, Trinh T (2010) The epidemiology, prevention, and detection of melanoma. Ochsner J 10(2):56–65
37. Sáez A, Sánchez-Monedero J, Gutiérrez PA, Hervás-Martínez C (2015) Machine learning methods for binary and multiclass classification of melanoma thickness from dermoscopic images. IEEE Trans Med Imaging 35(4):1036–1045
38. Sahlol AT, Yousri D, Ewees AA, Al-Qaness MA, Damasevicius R, Elaziz MA (2020) COVID-19 image classification using deep features and fractional-order marine predators algorithm. Sci Rep 10(1):1–15
39. Sharma N, Jain V, Mishra A (2018) An analysis of convolutional neural networks for image classification. Proced Comput Sci 132:377–384
40. Shorten C, Khoshgoftaar TM (2019) A survey on image data augmentation for deep learning. J Big Data 6(1):1–48. https://doi.org/10.1186/s40537-019-0197-0
41. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556
42. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016) Rethinking the inception architecture for computer vision. In proceedings of the IEEE conference on computer vision and pattern recognition, 2818-2826. https://doi.org/10.1109/CVPR.2016.308
43. Weese J, Lorenz C (2016) Four challenges in medical image analysis from an industrial perspective. Med Image Anal 33:44–49. https://doi.org/10.1016/j.media.2016.06.023
44. Weiss K, Khoshgoftaar TM, Wang D (2016) A survey of transfer learning. J Big Data 3(1):1–40. https://doi.org/10.1186/s40537-016-0043-6
45. Welch HG, Woloshin S, Schwartz LM (2005) Skin biopsy rates and incidence of melanoma: population based ecological study. BMJ 331(7515):481
46. Wu Y, He F, Zhang D, Li X (2015) Service-oriented feature-based data exchange for cloud-based design and manufacturing. IEEE Trans Serv Comput 11(2):341–353
47. Wu J, Hu W, Wen Y, Tu W, Liu X (2020) Skin lesion classification using densely connected convolutional networks with attention residual learning. Sensors 20(24):7080
48. Wu Z, Li S, Chen C, Hao A, Qin H (2020) A deeper look at salient object detection: bi-stream network with a small training dataset. arXiv preprint arXiv:2008.02938
49. Wu Z, Li S, Chen C, Hao A, Qin H (2022) Recursive multi-model complementary deep fusion for robust salient object detection via parallel sub-networks. Pattern Recogn 121:108212
50. Xie F, Fan H, Li Y, Jiang Z, Meng R, Bovik A (2016) Melanoma classification on dermoscopy images using a neural network ensemble model. IEEE Trans Med Imaging 36(3):849–858

51. Ying X (2019) An overview of overfitting and its solutions. J Phys Conf Ser 1168(2):022022 IOP Publishing
52. Yu L, Chen H, Dou Q, Qin J, Heng PA (2016) Automated melanoma recognition in dermoscopy images via very deep residual networks. IEEE Trans Med Imaging 36(4):994–1004
53. Zhang D, He F, Han S, Zou L, Wu Y, Chen Y (2017) An efficient approach to directly compute the exact Hausdorff distance for 3D point sets. Integr Comput-Aided Eng 24(3):261–277
54. Zhang J, Xie Y, Xia Y, Shen C (2019) Attention residual learning for skin lesion classification. IEEE Trans Med Imaging 38(9):2092–2103. https://doi.org/10.1109/tmi.2019.2893944