SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V11I8P131

Volume 11 Issue 8, 325-334, August 2024

Original Article

A Deep Ensemble Feature Extraction and Classification Framework for Melanoma Identification

Soumya Gadag[1], V. Panduranga Rao Malode[2]

[1]*Department of Electronics and Communication, S.G. Balekundri Institute of Technology, Karnataka, India.*
[1]*Department of Electronics Engineering, Jain Deemed to be University, Bangalore, Karnataka, India.*
[2]*Department of Computer Science Engineering, FET- JAIN Deemed to be University, Bangalore, India.*

[1]*Corresponding Author : soumya.gadag5@gmail.com*

Received: 27 June 2024 Revised: 04 August 2024 Accepted: 17 August 2024 Published: 31 August 2024

Abstract *- Skin cancer is a highly severe manifestation of the disease, with the potential to metastasize to other regions of the body if not identified in its early stages. According to WHO and the American Cancer Society (ACS), it is one of the leading causes of mortality in the worldwide population. Early discovery of skin cancer, however, can aid in a proper diagnosis to lessen the disease's effects on people. Dermoscopy Several methods have been introduced to develop automated Computer-Aided Diagnosis (CAD) systems where machine learning based solutions are widely adopted in this domain. To improve the detection accuracy, researchers have introduced deep learning-based solutions. However, the detection performance is affected by several factors, such as the uneven boundary of skin cancer, pigmentation, and hairs on dermoscopy images. Therefore, a novel approach for feature extraction by using handcrafted shape and texture features is introduced in this work. Moreover, the proposed approach adopts pre-trained deep learning architectures for deep feature extraction. The acquired features are combined and subjected to an ensemble classification technique, which uses decision trees, random forests, and support vector machines. A majority voting categorization is used to arrive at the final choice. The outcome of this approach is validated on publically accessible datasets such as PH2, ISIC 2017, and ISIC 2018.*

Keywords *- Classification, Deep Learning, Feature extraction, Machine Learning, Melanoma.*

1. Introduction

Recently, the world has noticed an increase in health-related issues in the worldwide population, where cancer has appeared among the most hazardous diseases, posing severe threats to human life. The term "cancer" is typically used to describe a wide spectrum of diseases that are all characterized by the body's abnormal cells proliferating quickly and out of control [1]. Several types of cancers have been identified as deadly, like cancer of the brain, prostate, breast, melanoma, and many more [2]. The WHO reports that one of the leading causes of death worldwide is cancer. Millions of individuals worldwide are impacted by skin cancer, making it a major threat to the general population. It is the most prevalent type of cancer, making up a sizable part of all cancer occurrences.

According to a study presented in [3], WHO has reported that around 2-3 million cases of non-melanoma are identified, whereas 132,000 cases of melanoma cancer were identified worldwide. Furthermore, according to the ACS, by the time they are 70 years old, one in five Americans is predicted to get skin cancer. In addition, the American Cancer Society (ACS) has revealed that melanoma is the most fatal kind of skin cancer, accounting for a sizable number of deaths. It is also the third most prevalent cancer among adults aged 20 to 39 [4]. The annual expense of treatment for these kinds of illnesses rises; for instance, skin cancer therapy in the US alone is estimated to cost over 3.3 billion USD per year. Patients diagnosed with early-stage melanoma have an approximate 98% estimated 5-year survival rate (American Cancer Society, 2020). According to a study, patients with Stage I melanoma who received treatment between 30 to 59 days after diagnosis had a 5% elevated risk of mortality against those who got treatment within 30 days.

On the other hand, patients with Stage I melanoma who underwent treatment more than 119 days after diagnosis faced a significantly higher risk, with a 41% increase in the likelihood of dying compared to those treated within the first 30 days [5]. Therefore, timely recognition of melanoma has a pivotal role in preserving human lives. The most common cause of skin cancer-related mortality is melanoma, the most aggressive and fatal kind of the disease. Despite its prevalence, the precise etiology of melanoma remains unclear [6]. However, numerous factors, such as genetic predisposition, exposure to ultraviolet radiation, and environmental influences, are implicated in the development of this condition.

Melanoma arises from the malignant transformation of skin melanocytes, which are responsible for producing dark pigments in various areas of the body, including the skin, hair, and eyes. Consequently, melanoma tumors predominantly exhibit brown or black hues. However, there are instances where melanomas lack pigment production, leading to their appearance as pink, red, or purple lesions [7]. As discussed before, the early detection of melanoma can be beneficial in preventing its deadly impact on human lives. Currently, image-processing-based solutions for melanoma detection have gained huge attention to develop the CAD system for skin lesion classification. For early recognition of skin cancer, dermoscopy is an effective non-invasive imaging technique widely employed by dermatologists [8]. This method magnifies the surface of the skin lesion, enhancing its structure and aiding the dermatologist in assessment.

Moreover, current advancement in machine learning is also adopted widely in computer vision-based applications like object detection, segmentation, and classification [9]. Several processes, including pre-processing, classification, segmentation, and feature extraction, are involved in typical machine learning techniques. The extraction of features from malignant images requires a substantial amount of knowledge. If segmentation is not performed adequately, it can lead to poor feature extraction, resulting in reduced classification accuracy. Similarly, robust feature extraction also plays an important role in achieving the desired accuracy. Several methods are derived for melanoma detection and classification utilizing traditional ML approaches. The authors published an SVM-based method for melanoma identification in [10]. However, the majority of these investigations used classifiers that were trained using manually created characteristics that were taken out of the dermoscopic pictures. Many ML techniques are time-consuming processes to achieve a precise diagnosis, and their efficiency relies heavily on the chosen attributes of the affected region on the skin lesion.

To overcome the issues of machine learning, deep learning-based systems have been adapted due to their significant capacity for pattern learning and achieving improved accuracy. In [11] presented a deep learning-based model for melanoma classification. An encoder-decoder module is used in this design to extract deep features. The pixel-wise classification module processes the features that have been gathered. To arrive at the final classification result, the lesion classification model is also used. Several deep learning based methods have been introduced, such as Lafraxo et al. [12] developed an automated process to detect and classify benign or malignant melanoma. Due to the current advancements of deep learning, authors have adopted DL based method and presented CNN-based architecture MelaNet. This architecture uses regularization, dropout and data augmentation processes to address the overfitting issue of CNN. Lu et al. [13] proposed CNN based architecture for skin cancer detection. This model considers XceptionNet as the base CNN architecture and incorporates the swish activation function with depthwise separable convolutions to introduce the improved XceptionNet. An automatic melanoma detection deep learning model was presented by Bhimavarapu et al. [14]. GrabCut-stacked CNN and a fuzzy basis are combined in this model for training. Further, an SVM classifier is used for classification which reported an overall accuracy of 98%. Banerjee et al. [15] presented DL based YOLO approach for melanoma detection. This method focuses on object detection and prediction of bounding boxes of detected objects along with its confidence score. Additionally, this approach employs a two-phase segmentation process that combines fuzzy logic-based approximations with graph theory. Islam et al. [16] adopted transfer learning with deep learning for skin lesion classification. DL methods achieve better performance when compared with traditional machine learning-based methods; however, computational complexity, robust classification performance, and handling the class imbalance problem remain challenging issues. Moreover, there is a scope to improve classification accuracy and reduce false positives. To address these problems, this article introduces a unique deep learning strategy built on top of pre-trained deep learning architecture. The proposed approach uses UNet based model to obtain the segmented region. Further, this segmented output is processed through the feature extraction module, where pre-trained deep learning architectures are used to extract the robust feature set. In the next stage, these features are fused to produce a fused feature vector. To determine the final classification result, a voting-based classifier model is finally presented.

1.1. Problem Statement

Melanoma, the deadliest form of skin cancer, accounts for a substantial number of cancer-related deaths, particularly among younger adults. Early detection of melanoma is essential for improving survival rates, yet it remains challenging due to its complex etiology involving genetic, environmental, and ultraviolet radiation factors. Despite advancements in medical imaging techniques like dermoscopy and machine learning-based diagnostic tools, achieving high accuracy in melanoma detection and classification remains difficult due to issues such as poor feature extraction and class imbalance. Therefore, there is an urgent need to develop more accurate, efficient, and robust methods for early melanoma detection and classification to reduce mortality rates and healthcare costs associated with skin cancer.

1.2. Research Gaps

Current methods for early melanoma detection, such as dermoscopy and traditional machine learning techniques, often fail to achieve high accuracy due to limitations in feature extraction and segmentation processes. Poor feature extraction can lead to reduced classification accuracy, highlighting the need for more effective methods. Similarly, existing machine learning approaches rely heavily on manually crafted features

and segmentation, which are time-consuming and may not always result in precise diagnosis. Although DL models have shown impressive results in melanoma detection, issues such as computational complexity, overfitting, and class imbalance remain significant challenges. There is a need for innovative deep learning strategies that address these issues effectively.

1.3. Contributions

The main contributions of this approach are as follows:

- First of all, a deep learning based UNet architecture for segmentation is presented to generate the segmented region for further analysis.
- After obtaining the region of interest, it is processed through the feature extraction step, when tasks related to texture and shape feature extraction are manually carried out.
- This approach also uses a deep feature process where three different pre-trained models, i.e. AlexNet, VGG16 and ResNet, are employed for deep feature extraction.
- After that, the acquired features are combined to create the final feature vector. The final feature vector is then fed to the ensemble classifier model, where three different classification models, namely Random forest, Decision tree and SVM, are used to formulate the ensemble model and a majority voting scheme is employed to obtain the final prediction.

Following is the arrangement of the remaining content in the article: Section 2 delivers a brief overview of existing schemes, section 3 offers a hybrid solution for robust feature extraction and ensemble classification, and results of the suggested approach are shown in section 4 along with a comparison to cutting-edge classification algorithms, and section 5 offers closing thoughts.

2. Literature Review

A synopsis of the current melanoma classification schemes is provided in this section. This work mainly focused on supervised machine learning based classification algorithms with different techniques of feature extraction. CNN-based UNet architecture was utilized by Seeja et al. [16] to segment images. Next, feature extraction methods comprise Gabor, Oriented Gradient Histogram, Local Binary Pattern (LBP), and Edge Histogram (EH). Finally, these feature vectors are fed into different classifiers such as SVM, RF, KNN, and Naïve Bayes to obtain the final classification. A method for automatic skin lesion detection via multilayer feature reduction and pixel-based seed segmentation was presented by Rehman et al. [17]. The complete method is divided into four different stages. Where the first stage performs contrast enhancement by using mean-based functioning, in the next stage, seed region growing and graph-cut based methods are introduced for segmentation, and segmented lesions are fused by applying the pixel-based fusion method to produce the final segmentation map. Further, HOG, color, and Speeded Up Robust Features (SURF) features are extracted and concatenated to formulate the final feature vector. To address the dimensionality issue, this method uses an entropy-based feature reduction approach, and the final feature vector is fed into the SVM classifier with a cubic kernel function. Tumpa et al. [18] reported that the traditional methods of melanoma detection rely on shape, color, and texture features of segmented images, but it does not guarantee to achieve the desired accuracy; therefore, to improve the performance, authors introduced border-line characteristics of lesion segmentation along with the combined gradients and LBP features. Later, these features are combined with the conventional features and are fused together and classified by using an SVM classifier. Bag-of-features-based systems for melanoma classification were covered by Hu et al. [19]. Codebook learning is a major task for these systems, and K-Means clustering is used to accomplish it. However, the systems have a problem with suboptimal codebook learning, which lowers overall classification accuracy. To address this problem, the scientists developed a new codebook learning approach that measures feature similarity to assess melanoma features more effectively. This scheme also considers the BoF fusion with different features such as color histogram and SIFT features. Hagerty et al. [20] presented a combined approach that uses a combined model of conventional image processing with the DL approach. The image processing module helps to detect dermoscopy information such as color distribution, blood vessels, and pigmentation. Further, it uses the ResNet-50 model for the transfer learning based classification model. İlkin et al. [21] developed a hybrid approach for melanoma classification, which considers SVM and heuristic optimization to formulate the combined classification. The Gaussian Radial Basis Function (RBF) learning module, which the SVM model uses, is improved further by the addition of the bacterial colony algorithm. To detect melanoma, Lee et al. [22] concentrated on segmentation, feature extraction, and classification. To accomplish the segmentation in this work, the authors used the edge-imfill method in conjunction with the Otsu method. Similarly, authors have suggested incorporating the ABC feature extraction with median thresholding.

3. Proposed Approach

This section presents a suggested DL-based method for melanoma detection feature extraction and machine learning classification. This model creates a robust feature vector by combining deep learning features, shape, and texture. The general architecture of the suggested model is shown in Figure 1, which is provided below. In this method, a UNet- is used to segment the dermoscopy images. The obtained segmented images are further given input to the feature extraction model, where the shape and texture features are extracted as handcrafted features. Similarly, the pre-trained AlexNet, VGG16, and ResNet Models are employed to extract the deep features [23]. In the next stage, these features are used to train

the multiple supervised classifiers like SVM, decision tree, and Random Forest [24]. To obtain robust classification performance, the majority voting scheme is applied to ensure the ensemble of these classification models. Finally, the outcome of this model is obtained as a multiclass classification.

3.1. Texture Feature Extraction

The proposed approach performs LBP feature extraction to obtain the texture features. One well-liked method in image processing and computer vision for characterizing the texture patterns in an image is the Local Binary Patterns (LBP) texture feature extraction [25] model. Let us assume a dermoscopy image is converted in grayscale represented as $I(x, y)$ where (x, y) denotes the pixel coordinates. The Local Binary Pattern (LBP) for this image can be computed as:

- Define a circular neighborhood with a radius R around each pixel (x, y). The neighbourhood will contain P sampling points evenly distributed on the circle.
- The center pixel value is denoted as $I_c = I(x, y)$.
- For each sampling point i ($i = 1\ to\ P$), its coordinates in the neighborhood are given by Equation (1) :

$$x_i = x + R * cos\left(2\pi * \frac{i}{P}\right)$$
$$y_i = y - R * cos\left(2\pi * \frac{i}{P}\right) \quad (1)$$

- Compare the intensity value of each sampling point with the center pixel value:
  If $I(x_i, y_i) >= I_c$, set the corresponding bit in the binary pattern to 1 (i.e., LBP(x, y, i) = 1).
  If $I(x_i, y_i) < I_c$, set the corresponding bit in the binary pattern to 0 (i.e., LBP(x, y, i) = 0).
- The LBP for the pixel (x, y) is represented as a binary number $LBP(x, y)$, which is created by concatenating the binary values from all P sampling points.
- After computing the LBP patterns for each pixel in the image, create a histogram H with K bins, where K is the total number of possible LBP patterns (usually 2^P for P-bit LBP).
- For each pixel (x, y) in the image, increment the corresponding bin in the histogram H based on its LBP value: $H[LBP(x, y)] = H[LBP(x, y)] + 1$
- The last feature vector, FV, signifies the texture information of the entire image. It is derived from the histogram H by concatenating the bin values: $FV = [H[0], H[1], H[2], \ldots, H[K-1]]$

3.2. Extraction Shape Feature Vector

Shape features play an imperative role in the identification of melanoma and its severity. This work extracts Hu moments as shape features, which is a set of seven invariant moments that capture the global shape information of an object. Let $I(x, y)$ represent the grayscale input image, where (x, y) are the pixel coordinates. The shape feature extraction performs raw central moment computation, computing the centroid coordinates central moments, normalizing the central moments, computing the Hu moments and achieving the rotation invariance.

- Compute the raw central moments as in Equation (2):

$$m_{pq} = \sum_x \sum_y x^p y^q I(x, y) \quad (2)$$

- Compute the centroid coordinates $\bar{x}$ and $\bar{y}$ as follows in Equation (3):

$$\bar{x} = \frac{m_{10}}{m_{00}}$$
$$\bar{y} = \frac{m_{01}}{m_{00}} \quad (3)$$

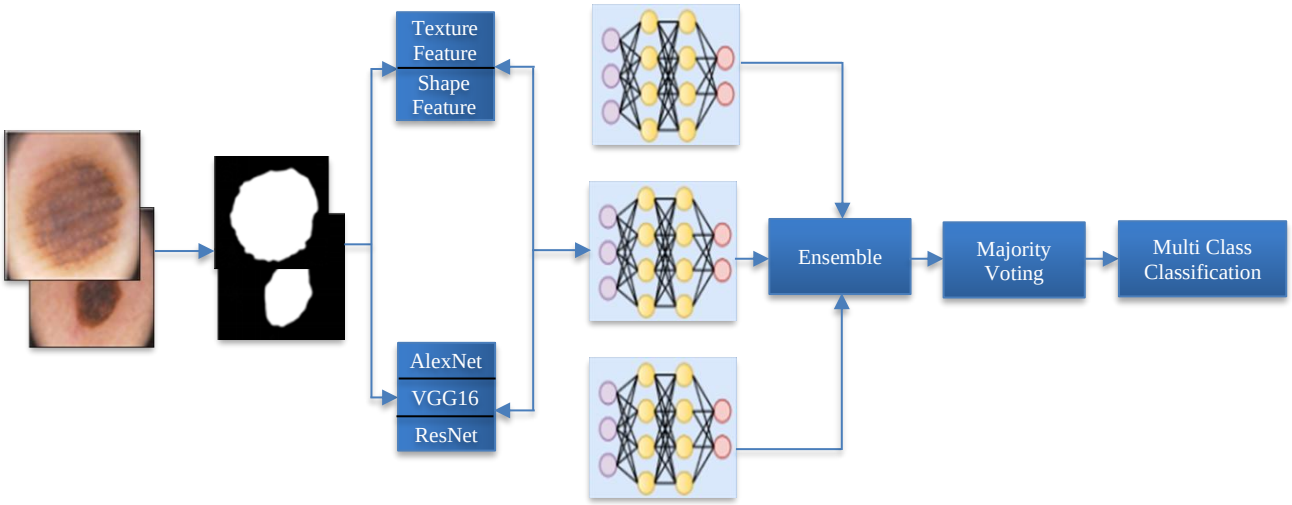

Fig. 1 Overall architecture of proposed DeepHand Feature extraction with ensemble approach

- Compute the updated central moments μ_{pq} as shown in Equation (4):

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q I(x, y) \quad (4)$$

- Perform central moment normalization as shown in Equation (5):

$$\mu_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\left(\frac{p+q}{2}+1\right)}} \quad (5)$$

- Compute the Hu moment I_{pq} as follows in Equation (6):

$$\begin{aligned} I_{20} &= \eta_{20} + \eta_{02} \\ I_{02} &= \eta_{20} + \eta_{02} \\ I_{11} &= \eta_{11} \\ I_{30} &= \eta_{30} + 3\eta_{12} \\ I_{12} &= \eta_{30} + 3\eta_{12} \\ I_{21} &= \eta_{21} + \eta_{03} \\ I_{03} &= \eta_{21} + \eta_{03} \end{aligned} \quad (6)$$

- Finally, normalize the obtained Hu moments to achieve the rotation invariance as the final shape feature vector, as shown in Equation (7):

$$I'_{pq} = sign\left(I_{pq}\right).\log_{10}\left(I_{pq}\right) \quad (7)$$

3.3. Deep Feature Vector

The application of pre-trained DL models for deep feature extraction is covered in this section. The suggested method makes use of VGGNet, ResNet, and AlexNet. These models are pretrained models and widely adopted for feature extraction tasks. The AlexNet helps to alleviate the vanishing gradient issue; similarly, VGGNet demonstrates the importance of deep networks for better feature learning while maintaining a simple and uniform structure, whereas the ResNet architecture enabled the training of extremely deep neural networks to accomplish complex computer vision tasks.

3.3.1. AlexNet [26]

AlexNet is a deep CNN architecture that played a pivotal role in the resurgence of interest in DL and its application to computer vision tasks. There are three fully linked layers and five convolutional layers totaling eight layers in it. By using pooling and convolution techniques, the first five layers are in charge of obtaining features. Initially, the network contains a convolution layer, and later max-pooling layer is placed, and the ReLU activation function with an 11x11 filter with a stride of 4 is also used. The following convolutional layers, which use smaller filter sizes (3x3) with a stride of 1, are again followed by max-pooling and ReLU activation. Additionally, it makes use of the ReLU activation function, which expedites training and helps address the vanishing gradient issue. Additionally, it uses Local Response Normalization (LRN) to normalize neuronal outputs across adjacent channels, improving the generalization capacity of the network. Finally, the dropout operation is applied to the fully connected layers to prevent overfitting during training. Figure 2 depicts the architecture of AlexNet used for deep feature extraction.

3.3.2. VGG16 [27]

It has sixteen layers total, of which three are fully linked layers, and the other thirteen are convolutional layers. Classification is performed using the fully connected layers, and max-pooling is performed after each block of convolutional layers. In order to maintain the spatial dimensions of the input feature maps, VGG16's convolutional layers all employ modest 3x3 filters with a stride of 1 and a padding of 1.

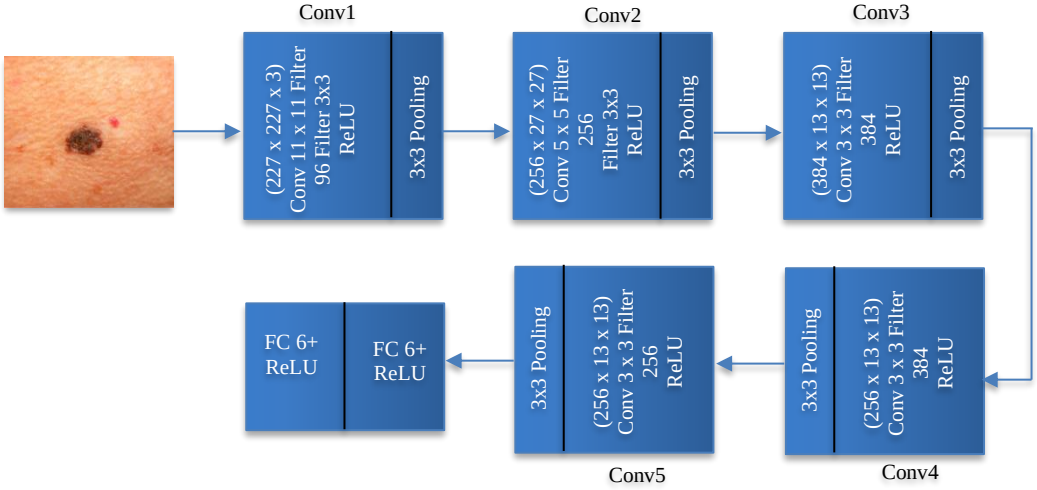

Fig. 2 AlexNet architecture

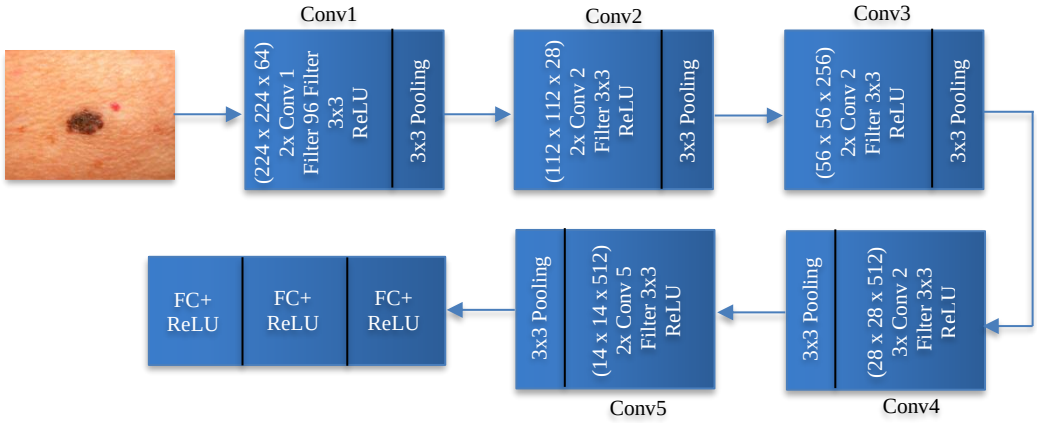

Fig. 3 VGG16 architecture

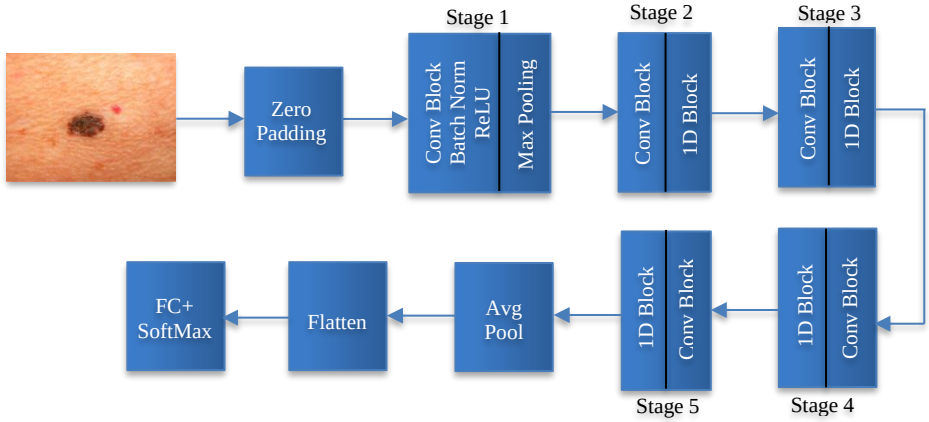

Fig. 4 ResNet architecture

The use of these small filters makes the network deeper while keeping the number of parameters relatively low. ReLU activation functions add non-linearity to the model by being applied after every convolutional layer. VGG16 uses a large number of filters in each layer (ranging from 64 to 512), making it computationally intensive but also allowing the model to learn more complex and abstract features. Figure 3 depicts the architecture of the VGG16 network.

3.3.3. ResNet [28]

It comprises 50 layers, including pooling, convolutional, fully connected, and shortcut connections (residual blocks) layers. It builds upon the original ResNet design, leveraging the power of residual learning to enable effective training of a much deeper neural network. This architecture starts with a traditional convolutional layer followed by four sets of residual blocks. Since each set consists of several residual blocks with progressively more filters, the network can learn more intricate features as it digs deeper. Utilizing 1x1, 3x3, and 1x1 convolutions together, it employs the bottleneck design. With minimal loss of expressive capability, this architecture drastically lowers the computational cost. The training of extremely deep networks is also made possible by the addition of skip connections between layers, which facilitate gradient flow during backpropagation and address the issue of vanishing gradients. At the end of the network, ResNet-50 replaces fully connected layers with global average pooling. This simplifies the model and decreases the parameter count, increasing its ability to generalize. ResNet-50 utilizes batch normalization, which normalizes the activations of each layer and helps stabilize and accelerate the training process. Given Figure 4 depicts the ResNet architecture.

3.4. Ensemble Classification Using Majority Voting

The acquired characteristics are then subjected to additional processing using the SVM, DT, and RF classic machine learning classification models. A majority vote process is used to determine the final classification decision. Several classifiers, or models, are combined in ensemble learning to arrive at a final decision. A straightforward and efficient classification approach employed in this situation is majority voting. The following is a description of the mathematical model for majority voting classification: Let us assume that there are N classifiers in total, represented by the notations C_1, C_2,... C_N. The same qualities that are acquired by

combining deep learning and handmade features are used to train each classifier. Furthermore, every classifier can categorize patterns into K distinct classes. The majority voting categorization can be used in the following ways for every given input sample X:

- Obtain each classifier's predictions for the input sample X.
- Determine how many times each class label appears in the set $\{P_1, P_2,... P_N\}$.
- Find the class label that appears the most frequently (the majority class). When two classes are tied, a predefined tie-breaking rule may be used.
- Assign the majority class label to the input sample X's final classification output. Equation (8) provides the following expression for it:

$$Final\ Classification = \arg\max_{k} \sum_{i=1}^{N} 1(P_i = k) \quad (8)$$

The ensemble model has several advantages over a single classification model, such as robustness to noise and outliers. Generally, the outliers and noisy samples have a severe impact on the performance of a single classifier; whereas the ensemble model helps to mitigate this impact by multiple decision mechanisms. Similarly, the ensemble models, especially methods like Random Forest and Bagging, build multiple base models and combine their predictions. This averaging or voting process helps in reducing overfitting, which can be a problem with complex single models.

Ensembles can handle class imbalances effectively. By combining the predictions from different models, ensembles can prevent biases towards the majority class, which is a common issue in imbalanced datasets. This approach uses Decision trees, SVM, and random forest classification models, which have different underlying principles and make different assumptions about the data.

Combining models with diverse characteristics can capture a wider range of patterns in the data, enhancing the ensemble's overall performance. Moreover, DT classifiers are good at capturing complex non-linear relationships, while SVMs are effective in handling high-dimensional spaces and capturing intricate decision boundaries.

Random forests, being an ensemble of decision trees, inherit their ability to model complex relationships. By combining these, the ensemble can handle a wide variety of data complexities.

4. Results and Discussion

This part describes the experimental examination of the suggested method and contrasts the results with other melanoma categorization strategies already in use. The first subsection provides an overview of the dataset, while the second subsection outlines the metrics that are used to assess the system's overall performance. The suggested ensemble deep-learning method for melanoma classification is compared in the last subsection.

4.1. Dataset Details

The outcome of the proposed methodology is assessed on a variety of datasets and compared the results with earlier schemes in this study. The three distinct datasets were used in this analysis which are PH2, ISIC 2017, and 2018 databases.

These datasets are available publically for melanoma detection, segmentation and classification purposes. Below given subsection describes these datasets briefly:

4.1.1. ISIC 2017 Dataset [29]

There are two thousand dermoscopic pictures in this collection, all of which were obtained by removing the surface reflection method. This method contributes to a more profound visualisation of the skin.

A human or semi-automated procedure has produced annotated ground truth labels for every image. The ISIC 2017 dataset can be accessed through the official ISIC Archive and is publicly available. The collection's high-quality images were created by the non-invasive imaging technique known as dermoscopy.

4.1.2. ISIC 2018 Dataset [30]

There are 2594 photos in this collection, and a ground truth mask is included for every image. This dataset defines three subtasks: classification, attribute identification, and lesion segmentation. The 520 photos in the test data component, 259 in the evaluation data subsection, and 1815 in the train data subsection comprise the three subsections of the dataset.

4.1.3. PH2 Dataset [31]

Classification images of various skin lesions, including nevi, seborrhoeic keratosis, and melanocytic lesions, are included in this dataset collection. These courses are fundamental to the training of machine learning models that are used to distinguish between benign and malignant skin diseases. There are 200 dermoscopic pictures of skin lesions in this collection.

4.2. Performance Measurement Parameters

The following subsection explains the many performance assessment metrics that are used to assert the proposed model's general efficacy. The confusion matrix, a representation of the total TP, FP, TN, and FN samples, is calculated to assess the effectiveness of the classification model.

The number of correctly predicted and incorrectly classified cases makes up the confusion matrix. Below given Table 1 demonstrates an example of a three-class confusion matrix.

Table 1. Representation of the confusion matrix

	Predicted Class			
	Class 1	**Class 2**	**Class 3**	**FN**
Actual Lesion Class	LC_{11}	LC_{12}	LC_{13}	$(LC_{12} + LC_{11})$
	LC_{21}	LC_{22}	LC_{23}	$(LC_{21} + LC_{23})$
	LC_{31}	LC_{32}	LC_{33}	$(LC_{31} + LC_{32})$
FP	$(LC_{21}+LC_{31})$	$(LC_{12}+LC_{32})$	$(LC_{13}+LC_{33})$	
TP	LC_{11}	LC_{22}	LC_{33}	

In a test set, a True Positive (TP) denotes an accurate prediction of the positive class by the classifier. True Negative (TN): This suggests that the classifier model predicts with accuracy the negative class from the provided test set. The true positive and true negative outcomes demonstrate the classifier's accuracy. These categories need to, nonetheless, correspond with the TP and TN values. When the classifier model estimates the positive class incorrectly, it results in a False Positive (FP).

False Negative (FN): Indicates that the classifier incorrectly identified the negative class. The confusion matrix facilitates the computation of various performance metrics for the proposed method, including total accuracy, precision, specificity, sensitivity, and F-measure. The ratio of accurate predictions to total predictions is used to determine accuracy, which is a measure of the rate of correct classifications. Computation of it is shown in Equation (9):

$$Accuracy = \frac{Tp+TN}{Tp+TN+Fp+FN} \quad (9)$$

Similarly, the specificity and sensitivity can be computed as:

$$specificity = \frac{TN}{TN+FP} \quad (10)$$

and,

$$senstivity = \frac{TP}{TP+FN} \quad (11)$$

Sensitivity (True Positive Rate) measures the proportion of actual positives that are correctly identified by the test whereas the Specificity (True Negative Rate): This measures the proportion of actual negatives that are correctly identified by the test.

4.3. Comparative Analysis

This section describes the obtained results of the suggested approach and compares it with existing machine and deep learning classification methods. Table 2 demonstrates the outcome of the PH2 dataset.

Table 2. Comparative performance for PH2 dataset

Method	Specificity	Sensitivity	Accuracy
SVM	-	97.90	98.20
Neural Network	92.50	98.10	97.00
AlexNet	92	97.50	94.8
VGG16	94.00	91.50	92.8
GoogleNet	97.00	94.50	95.8
InceptionV3	96.00	98.5	97.2
Proposed Model	97.80	98.90	98.5

According to the results above, the suggested method gets the best classification performance with respect to accuracy, specificity, and sensitivity, measuring 98.50%, 97.80%, and 98.90%, respectively. Subsequently, the effectiveness of the proposed model is measured using the ISIC 2017 dataset below illustrated Table 3 demonstrates the comparison analysis.

Table 3. Comparative analysis for ISIC 2017 dataset

Methods	Specificity	Sensitivity	Accuracy
MLP	72.16	93.20	87.88
SVM	89.52	93.81	92.72
KNN	86.52	94.42	92.42
Inception ResNet	95.51	96.45	96.21
Proposed Model	97.20	98.10	98.55

A variety of classifiers, including KNN, MLP, SVM, and Inception ResNet, are used to compare the resulting performance. Additionally, the ISIC 2019 dataset performance was assessed, and the results are shown in Table 4.

Table 4. Comparative performance for ISIC 2018 dataset

Methods	Specificity	Sensitivity	Accuracy
KNN	86.43	93.13	91.39
MLP	80.66	91.25	88.50
SVM	85.33	92.73	90.81
Inception ResNet	94.67	96.67	96.31
Proposed Model	98.10	97.90	98.40

This experiment demonstrates that the projected scheme attains better classification performance for the ISIC 2018 dataset by achieving average performance of 98.40%, 98.10%, and 97.90% in terms of Accuracy, Specificity, and Sensitivity.

The KNN can be computationally expensive because it requires computing the distance to all training samples for each prediction, and it is sensitive to noisy data. Similarly, the computational complexity affects the performance of SVM because of quadratic optimization problems, and the Inception ResNet model suffers from overfitting issues. The proposed combination helps to overcome these drawbacks by presenting simple deep learning based architecture.

5. Conclusion and Future Work

Skin melanoma classification is an essential task in dermatology and can significantly impact patient outcomes. In recent years, two prominent approaches for extracting features in this context have been handcrafted features and pre-trained deep learning features. Handcrafted features offer interpretability and ease of implementation. By utilizing domain knowledge and expertise, specific texture, color, and shape features can be designed to capture relevant information for melanoma classification. These features are generally lightweight and computationally efficient, making them suitable for resource-constrained environments. They could, however, have trouble identifying the intricate changes and patterns shown in melanoma photos, which would reduce the precision of their categorization. However, pre-trained deep learning features especially those derived from Convolutional Neural Networks (CNNs) have demonstrated impressive performance across a range of image classification tasks, including skin melanoma classification.

Pre-trained CNN features can better distinguish between benign and malignant skin lesions by utilizing the hierarchical representations that have been acquired from large-scale image datasets to capture high-level abstract information. These features excel in handling complex structures and variations, leading to higher classification accuracy. This work presents a combined approach by using these feature extraction techniques.

The obtained features are classified with the help of an ensemble classification approach by using a majority voting scheme. Accuracy, sensitivity, and specificity are the metrics used to quantify the performance of this technique, which is tested on publicly accessible datasets. The outcome shows that the suggested strategy works better than modern approaches to deep learning and machine learning.

References

[1] Sunil Kumar et al., "A Chaos Study of Tumor and Effector Cells in Fractional Tumor-Immune Model for Cancer Treatment," *Chaos, Solitons & Fractals*, vol. 141, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[2] Rebecca L. Siegel et al., "Cancer Statistics, 2021," *CA: A Cancer Journal for Clinicians*, vol. 71, no. 1, pp. 7-33, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[3] Faouzi Adjed et al., "Classification of Skin Cancer Images Using Local Binary Pattern and SVM Classifier," *AIP Conference Proceedings*, vol. 1787, no. 1, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[4] Pamela Hallquist Viale, "The American Cancer Society's Facts & Figures: 2020 Edition," *Journal of the Advanced Practitioner in Oncology*, vol. 11, no. 2, pp. 135-136, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[5] Melina Tziomaka, and Ilias Maglogiannis, "Ensembles of Deep Convolutional Neural Networks for Detecting Melanoma in Dermoscopy Images," *Computational Collective Intelligence: 13th International Conference*, Rhodes, Greece, pp. 523-535, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[6] Ni Zhang et al., "Skin Cancer Diagnosis Based on Optimized Convolutional Neural Network," *Artificial Intelligence in Medicine*, vol. 102, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[7] Ahmed Alsayyah, "Differentiating between Early Melanomas and Melanocytic Nevi: A State-of-the-Art Review," *Pathology - Research and Practice*, vol. 249, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[8] Saqib Ali et al., "State-of-the-Art Challenges and Perspectives in Multi-Organ Cancer Diagnosis via Deep Learning-Based Methods," *Cancers*, vol. 13, no. 21, pp. 1-23, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[9] Anamika Dhillon, and Gyanendra K. Verma, "Convolutional Neural Network: A Review of Models, Methodologies and Applications to Object Detection," *Progress in Artificial Intelligence*, vol. 9, pp. 85-112, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[10] Praveen Banasode, Minal Patil, and Nikhil Ammanagi, "A Melanoma Skin Cancer Detection Using Machine Learning Technique: Support Vector Machine," *IOP Conference Series: Materials Science and Engineering: 1st International Conference on Frontiers in Engineering Science and Technology*, Mangalore, India, vol. 1065, pp. 1-5, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[11] Adekanmi A. Adegun, and Serestina Viriri, "Deep Learning-Based System for Automatic Melanoma Detection," *IEEE Access*, vol. 8, pp. 7160-7172, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[12] Samira Lafraxo, Mohamed El Ansari, and Said Charfi, "MelaNet: An Effective Deep Learning Framework for Melanoma Detection Using Dermoscopic Images," *Multimedia Tools and Applications*, vol. 81, pp. 16021-16045, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[13] Xinrong Lu, and Y.A. Firoozeh Abolhasani Zadeh, "Deep Learning-Based Classification for Melanoma Detection Using XceptionNet," *Journal of Healthcare Engineering*, vol. 2022, no. 1, pp. 1-10, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[14] Usharani Bhimavarapu, and Gopi Battineni, "Skin Lesion Analysis for Melanoma Detection Using the Novel Deep Learning Model Fuzzy GC-SCNN," *Healthcare*, vol. 10, no. 5, pp. 1-14, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[15] Shubhendu Banerjee et al., "Melanoma Diagnosis Using Deep Learning and Fuzzy Logic," *Diagnostics*, vol. 10, no. 8, pp. 1-26, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[16] R.D. Seeja, and A. Suresh, "Deep Learning Based Skin Lesion Segmentation and Classification of Melanoma Using Support Vector Machine (SVM)," *Asian Pacific Journal of Cancer Prevention*, vol. 20, no. 5, pp. 1555-1561, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[17] Amjad Rehman et al., "Microscopic Melanoma Detection and Classification: A Framework of Pixel-Based Fusion and Multilevel Features Reduction," *Microscopy Research and Technique*, vol. 83, no. 4, pp. 410-423, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[18] Priyanti Paul Tumpa, and Md Ahasan Kabir, "An Artificial Neural Network Based Detection and Classification of Melanoma Skin Cancer Using Hybrid Texture Features," *Sensors International*, vol. 2, pp. 1-8, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[19] Kai Hu et al., "Classification of Melanoma Based on Feature Similarity Measurement for Codebook Learning in the Bag-of-Features Model," *Biomedical Signal Processing and Control*, vol. 51, pp. 200-209, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[20] Jason R. Hagerty et al., "Deep Learning and Handcrafted Method Fusion: Higher Diagnostic Accuracy for Melanoma Dermoscopy Images," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 4, pp. 1385-1391, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[21] Sümeyya İlkin et al., "hybSVM: Bacterial Colony Optimization Algorithm Based SVM for Malignant Melanoma Detection," *Engineering Science and Technology, an International Journal*, vol. 24, no. 5, pp. 1059-1071, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[22] Hyunju Lee, and Kiwoon Kwon, "Diagnostic Techniques for Improved Segmentation, Feature Extraction, and Classification of Malignant Melanoma," *Biomedical Engineering Letters*, vol. 10, pp. 171-179, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[23] Khairul Islam et al., "Melanoma Skin Lesions Classification Using Deep Convolutional Neural Network with Transfer Learning," *2021 1st International Conference on Artificial Intelligence and Data Analytics*, Riyadh, Saudi Arabia, pp. 48-53, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[24] Amer Sallam et al., "Early Detection of Glaucoma Using Transfer Learning from Pre-Trained CNN Models," *2021 International Conference of Technology, Science and Administration*, Taiz, Yemen, pp. 1-5, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[25] Muhammad Zain Amin, and Amir Ali, "Performance Evaluation of Supervised Machine Learning Classifiers for Predicting Healthcare Operational Decisions," *Wavy AI Research Foundation*, pp. 1-7, 2018. [Google Scholar]

[26] Siti Salbiah Samsudin et al., "Skin Lesion Classification Using Multi-Resolution Empirical Mode Decomposition and Local Binary Pattern," *Plos One*, vol. 17, no. 9, pp. 1-14, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[27] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84-90, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[28] Karen Simonyan, and Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv*, pp. 1-14, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[29] Brett Koonce, *ResNet 50*, Convolutional Neural Networks with Swift for Tensorflow, Apress, Berkeley, CA, pp. 63-72, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[30] Noel C.F. Codella et al., "Skin Lesion Analysis Toward Melanoma Detection: A Challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), Hosted by the International Skin Imaging Collaboration," *2018 IEEE 15th International Symposium on Biomedical Imaging*, Washington, DC, USA, pp. 168-172, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[31] PH² Database, Automatic Computer-Based Diagnosis System for Dermoscopy Images, 2022. [Online]. Available: https://www.fc.up.pt/addi/ph2%20database.html