

RESEARCH

Open Access



Skin cancer classification using non-local means denoising and sparse dictionary learning based CNN

Apeksha Pandey¹, Manepalli Sai Teja¹, Parul Sahare^{1*} , Vipin Kamble², Mayur Parate¹ and Mohammad Farukh Hashmi³

*Correspondence:
parulsahare2387@gmail.com

¹ Department of Electronics and Communication Engineering, Indian Institute of Information Technology Nagpur, Nagpur, Maharashtra 441108, India

² Department of Electronics and Communication Engineering, Visvesvaraya National Institute of Technology, Nagpur, Maharashtra 440010, India

³ Department of Electronics and Communication Engineering, National Institute of Technology Warangal, Hanamkonda, Telangana 506004, India

Abstract

Skin conditions are becoming increasingly prevalent across the world in current times. With the rise in dermatological disorders, there is a need for computerized techniques that are completely noninvasive to patients' skin. As a result, deep learning models have become standard for the computerized detection of skin diseases. The performance efficiency of these models improves with access to more data with their primary aim being image classification. In this paper, we present a skin disease detection methodology using image processing techniques, non-local means denoising and convolutional neural network (CNN) backed by sparse dictionary learning. Here, the major benefit of using NLM denoising followed by sparse dictionary learning with CNNs in image classification lies in leveraging a multi-stage approach that enhances the quality of input data, extracts meaningful and discriminative features, and improves the overall performance of the classification model. This combined approach addresses challenges such as noise robustness, feature extraction, and classification accuracy, making it particularly effective in complex image analysis tasks. For denoising, the average Peak Signal to Noise Ratio (PSNR) obtained for images from HAM-10000 dataset is 33.59 dB. For the ISIC-2019 dataset, the average PSNR for the train folder is 34.37 dB, and for the test folder it is 34.39 dB. The deep learning network is trained for the analysis of skin cancer images using a CNN model and is achieving acceptable results in classifying skin cancer types. The datasets used contain high-resolution images. After all the tests, the accuracy obtained is 85.61% for the HAM-10000 dataset and 81.23% for the ISIC-2019 dataset, which is on par with existing approaches validated by benchmarking results.

Keywords: Skin disease, Non-local means denoising, Sparse dictionary learning, Skin cancer classification, Convolutional neural network

Introduction

The skin plays a key role by providing a barrier against invading pathogens and protecting against bacterial, fungal, and parasitic infections, making it a very important organ. An aberrant state of the skin is termed a skin disease. Therefore, the proper diagnosis of skin diseases has a crucial role in determining the most effective treatment plan and achieving successful outcomes at the earliest.

Skin disease such as skin cancer, if diagnosed and treated early, have a substantial possibility of recovery. The World Health Organization (WHO) has stressed on the fact that the prevalence of skin diseases in India is estimated to account for 10–12% of the population [1]. There are very few skilled dermatologists available to attend to the vast population. Experienced dermatologists conventionally follow a procedure, initially conducting naked-eye detection of suspected lesions, followed by dermatoscopy, and then performing a biopsy to confirm the condition. This whole process is invasive to the patient's skin, takes a significant amount of time, and the individual might advance to a subsequent stage [2]. Furthermore, reliable diagnosis is eccentric, depending on the practitioner's expertise. These factors make the manual detection of skin diseases one of the challenging and onerous tasks for patients. The prognosis may also vary for a particular condition with the proficiency of clinicians and their familiarity with different dermatoscopy procedures being used in which they have received professional and formal training. Hence, for the purpose of early phase diagnosing of skin cancer and addressing the related issues, thorough exploration has been conducted to enhance algorithms for computer-based image analysis.

There are various traditional or handcrafted feature-based classifications where the feature selection process is time-consuming and the selection of pertinent features is critical for the correct detection. There are several machine learning models such as SVM and logistic regression that can be utilized for classification. The other efficient way of tackling disease classification is by employing Deep Learning models. Deep Learning is used widely and extensively as a classification method because it helps in calculating features with ease within the convolutional layers and has been proven highly efficient outperforming classical methods providing much better accuracy. For Convolutional Neural Networks (CNNs), various pre-trained models like U-Net, ResNet, AlexNet and MobileNet have been developed and their efficiency is constantly evolving. The key to CNN's success is that the feature extraction part is not fixed. It provides a framework that is flexible enough to learn a suitable feature representation for the given task and yet constrained enough to be tractable and avoid the issue of overfitting. The advantage of the classifier component is its convenience, as it facilitates joint training through backpropagation, while the learned feature extraction component of the network remains intact.

In this paper, image processing and CNN along with sparse dictionary learning are employed for the detection of dermatological disorders in humans. The paper underlines the present techniques utilized for the detection of skin diseases, presents an approach for detection, and lists the numerous benefits. The article also includes an in-depth exploration of the various transformations employed for the implementation of the proposed method and the comparisons made are based on the accuracy parameter.

The main propositions for this paper are:

- Proposal of a CNN Model: This paper aims to propose a CNN model approach to aid in the development of digital models for detecting skin diseases.
- Image Processing and Noise Removal: To significantly improve accuracy, the use of Non-local Means for denoising is proposed.

- **Sparse Dictionary Learning:** Implementation of sparse dictionary learning to further enhance the classification efficiency of the model. Non-Local Means (NLM) denoising excels in preserving structural details and improving image quality by averaging similar patches from across the image, rather than just nearby pixels. This method effectively reduces noise by leveraging the redundancy of natural image textures and structures. NLM's adaptive filtering, which adjusts based on local content, allows it to handle varying noise levels and textures robustly. The result is enhanced perceptual quality, maintaining sharpness and clarity without excessive blurring. Its versatility makes it ideal for medical imaging applications, where high image quality is essential. Sparse-based dictionary learning provides a compact and efficient data representation using a sparse linear combination of atoms from a learned dictionary. This method requires only a few atoms to approximate each data point, resulting in a more compact representation than other techniques. It reduces storage needs and speeds up computational tasks in classification and image processing. The learned dictionary captures the underlying data structure, making it an effective feature extraction tool that uncovers important patterns. Additionally, sparse representations are robust to noise, as they focus on significant coefficients, minimizing the impact of noise on data representation. The dictionary adapts to the specific features of data, further enhancing its ability to capture domain-specific features, making sparse-based dictionary learning efficient and effective for data representation. The datasets utilized in this work include:
 - **HAM-10000 Dataset:** An open-source dataset available on the internet for public use, containing 10,015 images of seven types of pigmented skin lesion images.
 - **ISIC-2019 Dataset:** A dataset with a similar classification of pigmented skin lesion images, also used for this study.

Further, the framework for this paper is as follows: The previous work is mentioned in Section “[Literature survey](#)”. In Section “[Proposed method](#)”, proposed work is explained in detail. Experiments and results are discussed in Section “[Experiments and results](#)”. Finally, conclusion is given in Section “[Conclusion and future scope](#)” with future scope.

Literature survey

For the rapid detection and classification of skin diseases and to address the difficulties in diagnosing these conditions, many research solutions have been devised over the years by improving computer image analysis algorithms. Although, the type of data cannot be regulated, the algorithms have potential for improvement, and can more precisely diagnose diseases. Consequently, many researchers are continuously working on developing efficient systems, and some relevant work is discussed below.

In the work by [3] novel image processing techniques and geometry based features of the skin cancer involving ABCD and geometrical features of lesions are used for the detection of melanoma lesions. In [4], the authors aimed to classify the skin diseases by preprocessing, segmentation and feature extraction using the ABCD rule which assesses parameters such as asymmetry, edge irregularity, color, and lesion diameter. They achieved accuracy of up to 91.6% for detecting melanoma.

As discussed in [5], the authors focused on the detecting melanocytic lesions using a dataset collected from Dermweb, where segmentation is executed using clustering and SVM for classification. The accuracy obtained is 96.8%. However, this study emphasizes only on the detection of melanocytic skin lesions, i.e. melanoma. In similar work, presented in [6], a pre-trained neural network model is used for feature extraction and Error-Correcting Output Codes (ECOC) SVM classifier is employed for skin cancer classification [7] demonstrated that while CNN alone achieved an accuracy of about 91%, it was enhanced to 95.3% when combined with SVM. In [8], an SVM classifier is used with extracted features of shape, color and GLCM, to classify images into malignant or benign classes. Another study on the SVM classifier is discussed in [9] where color segmentation combined with an SVM classifier recognized eight classes of skin diseases with a performance of 94.79%.

The authors of [10] applied K-means clustering to gamma-corrected images for melanoma segmentation using GLCM to compute textural features from segmented images. Their model distinguishes melanoma into various types, such as nodular, acral, superficial and lentigo achieving 90% accuracy.

In [11] the authors used the Gaussian Radial Basis Kernel in SVM where, segmentation was performed by the GrabCut algorithm. Features like shape, color and geometry were computed using image processing techniques which were further classified as non-cancerous (mole benign) or cancerous (malignant). The best feature combination achieved an accuracy of 86.67%.

In [12], a method that operates on color image input precised and passed to a trained CNN achieved a 100% performance rate for detecting three different types of skin diseases. The features were classified using SVM, with results presented to the user including details on disease type, severity and extent of spread.

In [13], a dual-stage approach combining Computer Vision and Machine Learning was proposed. It uses maximum entropy with ANN in the first stage, followed by KNN, decision trees, and ANN for prediction in the second stage, achieving an accuracy of up to 95% for six diseases. The authors in [14] also proposed a system for detecting nine types of dermatological skin conditions. They first pre-processed color skin images to extract significant features followed by disease identification using ANN, in a two phase technique, achieving an accuracy rate of 90%. According to [15], their study focuses on the developing a malignant melanoma detection system using dermatoscopic images and available patient records. The model, which works on resource-poor devices, provides a balanced CNN+ANN model with higher accuracy of (92.34%) as compared to the CNN model (73.69%), based on the ISIC Archive dataset.

As outlined by [16], their proposed model aggregates a robust CNN into a framework and performs final classification based on the weighted outputs of member CNNs.

In [2], the authors employed a deep learning model, to analyze cancer images from more than 24,000 of skin samples in the benchmarked ISIC dataset using three ConvNet architectures: VGG-19, ResNet and InceptionV3. Numerous parameters were utilized to determine the optimal architectures for classification. The best accuracy of 86.90% was achieved with InceptionV3, which also yielded a precision of 87.47%, sensitivity of 86.14%, and specificity of 87.66%. Similarly, the image classification task in [17] used Inception V3 and MobileNet V1. Here, Inception V3 model achieved

72% accuracy while MobileNet V1 showed 58% accuracy. Research in [18] attained an overall accuracy of 83.1% using a pre-trained MobileNet.

In recent work, [19], the authors developed a deep learning framework using CNNs to classify skin diseases from dermoscopy images, using the CNN model, specifically VGG-16 to generalize and accurately predict skin diseases. Finally, the predictions are analyzed and the results are displayed through a web application, providing a practical interface for users to view the outcomes. Another recent work on VGG, [20], investigates the application of machine learning (ML) algorithms, particularly CNNs, in diagnosing skin diseases using the ISIC and Dermofit datasets, which contain 26,150 images. The study found that the CNN model, especially with Transfer Learning (TL) using VGG19, achieved a high accuracy of 94.91% in detecting various skin conditions as measured by metrics like Receiver Operating Characteristics (ROC) and Mean Squared Error (MSE).

In another work [21], the authors focused on the classification of nine skin diseases that share similar physical characteristics using a diverse dataset, preprocessing techniques were employed to prepare the data for training and testing. The LeNet architecture was then applied to classify these diseases, achieving a 95% accuracy rate. The proposed deep learning model not only demonstrated state-of-the-art performance in skin disease detection but also excelled in segmentation tasks, accurately delineating diseased regions. The model achieved a sensitivity of 90%, specificity of 94%, precision of 92%, and recall of 94%, underscoring its effectiveness in skin disease diagnosis.

In the work proposed by [22], introduced a new deep learning model based on optimal probability, where features extracted from pre-processed images were applied to the model for training using MATLAB software. This approach achieved an accuracy of 95%, specificity of 0.97, and sensitivity of 0.91. According to [23], a region-based CNN combined with fuzzy based k-means clustering was employed. In this method, visual information underwent enhancement before applying the faster RCNN to obtain fixed-length feature vectors. FKM was then used for segmenting skin regions affected by melanoma. This method delivered an average accuracy of 95.40% on the ISIC-2016 dataset, 93.1% on ISIC-2017, and 95.6% on the PH2 dataset.

The work by [24] integrated imaging modalities and patient metadata to improve the precision of automated skin lesion diagnosis. They used a modified ResNet-50 architecture and a multimodal classifier, surpassing the capabilities of a baseline classifier for melanoma detection.

In [25], a DCNN with a cross-modality learning strategy was proposed, extracting comprehensive features from sub-networks using Class Activation Mapping-Bilinear Pooling (CAM-BP). This technique produced probability maps that improved overall performance and facilitated the decision-making.

The authors in [26], worked on a modified GoogLeNet algorithm, to reduce the time complexity and achieved the highest classification accuracy of 0.9309.

In [27], the authors proposed an automated system for computer-aided diagnosis that focused on the classification of multi-class skin (MCS) cancer. The fine-tuning was performed on the seven classes from the HAM-10000 dataset, comparing five trained CNNs and four ensemble models. They reported a maximum accuracy of 93.20% with the individual ResNeXt101 model and 92.83% with the ensemble model.

ResNeXt101 was recommended for the MCS cancer classification due to its optimally designed architecture and superior accuracy.

In [28], an acne type classification system using a custom CNN sequential model was studied, examining accuracy variations based on model parameters. The level of accuracy ranged from 90%-95%, with skin sensitivity and acne density varying from 93 to 96%. [29], proposed a custom CNN model, with picture handling strategies on a dormant dataset of 500 images, successfully detecting various diseases with the precision of 73%. [30] employed a CNN based algorithm trained on clinical images encompassing forty skin diseases. This customized variant of Densenet-161 with a hybrid implementation combining focal loss and log loss, achieved $76.93 \pm 0.88\%$ accuracy and average 0.95 ± 0.02 AUC on a collection of affected skin images. The clinical diagnosis of a broad range of skin conditions highlights the potential of smartphone applications powered by artificial intelligence as a convenient point-of-care guidance tools.

In recent work by [31], a hybrid architecture combining deep CNN techniques inspired by pre-trained models was proposed for detection of skin cancer. It included three core techniques, uniform distribution of convolutional filters throughout the entire architecture, use of residual connections to address vanishing gradient and a cyclic learning rate for annealing. Authors of [32], described a model, that presented a cascading of ensemble networks that leverages an integrated ConvNet and multi-layer perceptron based on computed handcrafted features. This approach has shown significant improvement in the performance of ensemble model, achieving increase from 85.3 to 98.3%, compared to CNN.

In [33], comprehensive survey of various deep learning techniques for diagnosing skin diseases was presented, including the performance evaluation metrics for various deep learning architectures, popular frameworks, and algorithms.

Hence, the reviewed literature in this paper demonstrated that machine learning in skin disease detection is a rapidly growing field with tremendous potential to enhance diagnostic accuracy and efficiency, potentially revolutionizing the field of dermatology.

Proposed method

The proposed method involves a multi-step process, beginning with pre-processing, which includes denoising the image, followed by the implementation of sparse dictionary learning to achieve enhanced accuracy, and finally using a CNN model for classification. The block diagram for the same is demonstrated in Fig. 1.

Image preprocessing

The initial stage of the process involves performing resizing, denoising, and data augmentation, which are critical for preparing the data for accurate and efficient analysis. The steps for the same are portrayed in Fig. 2.

The starting point is to undertake resizing, where the images are resized to 100×100 . Further, the image noises are removed using Non-local means Denoising algorithm which is done using OpenCV library.

Non-Local Means (NLM) denoising is being used for denoising, due to its excellent capability to preserve image details and textures while effectively handling various types of noise. A comparative study for some denoising techniques is provided in

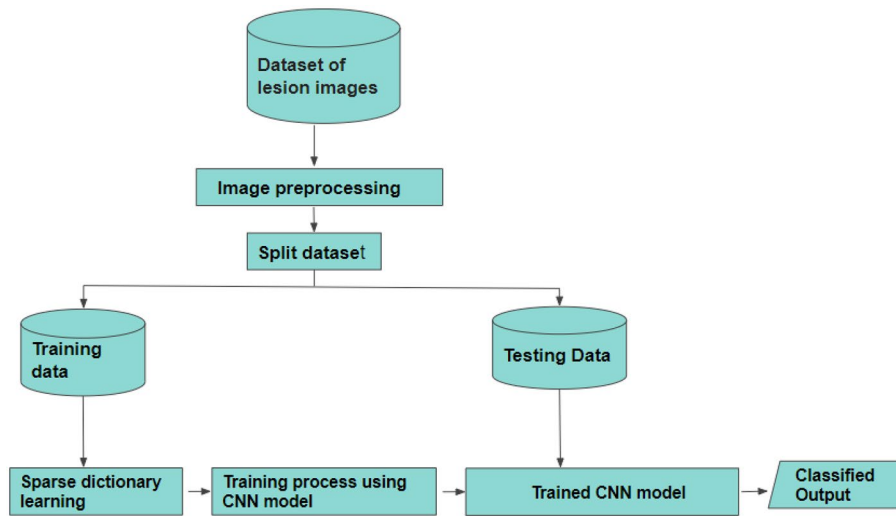


Fig. 1 Block diagram of the proposed model

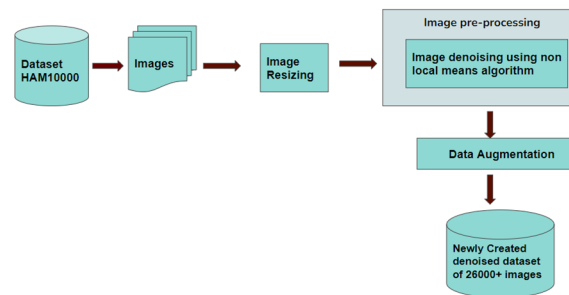


Fig. 2 Preprocessing and data augmentation

Table 1 The Non-local Means Denoising works efficiently for both Gaussian as well as Salt and Pepper noise, and also preserves skin textures better than other techniques. This algorithm works on fundamental principle that the colors of multiple image sub-windows which are similar to the pixel neighborhood can be averaged to replace the color of a single pixel. From Fig. 3, it is seen that in the denoising algorithm, pixel value is substituted by computing the mean from the set of neighboring pixels. This process involves the comparison of the patches which are centered on the neighboring pixels with the patch, centered on the pixel of interest. Here, only those pixels which have patches that are similar to the current patch are used for the averaging purpose. The computation of non-local means is given in Eq. 1.

$$NLu(p) = \frac{1}{C(p)} \cdot \int f(d(B(p), B(q)))u(q)dq \tag{1}$$

where $N Lu(p)$ represents the result of applying the Non-Local Means filter at pixel p , $d(B(p), B(q))$ denotes Euclidean distance of the image patches centered respectively at p and q . $u(q)$ is the pixel value at q , which is weighted by the function f based on the distance between patches. $C(p)$ denotes normalization and f is a decreasing function.

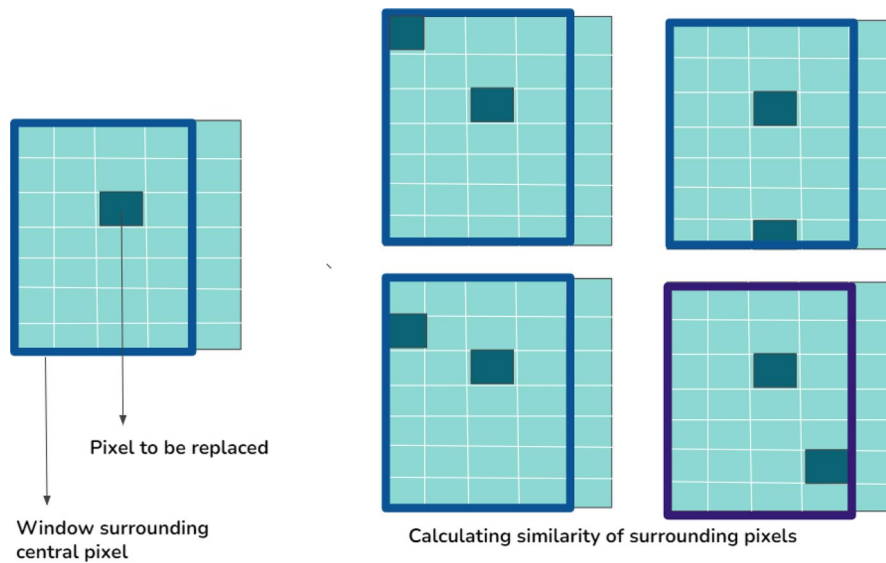


Fig. 3 Non-local means denoising

Table 1 Comparison of denoising techniques for skin cancer image preprocessing

Denoising technique	Advantages	Disadvantages compared to NLM
Gaussian blur	<ul style="list-style-type: none"> Simple and fast implementation Effective for removing Gaussian noise Good for basic preprocessing 	<ul style="list-style-type: none"> Blurs image edges May cause loss of fine details and textures Less effective for preserving structure in skin images
Median filtering	<ul style="list-style-type: none"> Excellent for removing salt-and-pepper noise Preserves edges better than Gaussian blur Simple and computationally efficient 	<ul style="list-style-type: none"> May not effectively handle Gaussian noise Can introduce blockiness or artifacts in homogeneous regions Less effective for preserving intricate details in skin textures
Wavelet thresholding	<ul style="list-style-type: none"> Good at preserving image features while reducing noise Can adapt to different noise levels Effective for multi-scale analysis 	<ul style="list-style-type: none"> Requires selection of appropriate wavelet and threshold Computationally more complex May not perform as well with non-stationary textures in skin images

For pixel-wise implementation,

$$ui(p) = \frac{1}{C(p)} \sum_{q \in B(p,r)} ui(q)w(p,q) \tag{2}$$

$$C(p) = \sum_{q \in B(p,r)} w(p,q) \tag{3}$$

where $ui(p)$ represents the denoised value at pixel p , $C(p)$ is the normalization factor for pixel p . $ui(q)$ is the pixel value at q , which is weighted by $w(p,q)$. $w(p,q)$ is the weight assigned to pixel q based on the similarity between the patch centered at p and the patch centered at q .

Further for a patch this could be represented as,

$$Bi = \frac{1}{C} \sum_{Q=Q(q,f) \in B(p,r)} ui(Q)w(B, Q) \tag{4}$$

$$C = \sum_{Q=Q(q,f) \in B(p,r)} w(B, Q) \tag{5}$$

$$ui(p) = \frac{1}{N^2} \sum_{Q=Q(q,f) | q \in B(p,f)} Qi(p) \tag{6}$$

where Bi represents the denoised value of the patch centered at pixel p , C is the normalization factor for patch B . $ui(Q)$ is the value of the patch Q . $w(B, Q)$ is the weight assigned to patch Q based on its similarity to patch B . N is normalization constant, representing the size of the neighborhood.

The implementation has a patch size of 7×7 and a window size of 21×21 . This algorithm is capable of effectively restoring textures that have been blurred by other denoising algorithms. Figure 4 illustrates the flowchart describing the main steps of the denoising algorithm. In the proposed work, the mean of *Peak Signal to Noise Ratio* i.e. *PSNR* obtained for 10,015 images in HAM-10000 is 33.59 dB. In the similar manner for ISIC dataset, the average values obtained are 34.37 dB and 34.39 dB for train and test folders respectively which are admissible values for effective denoising.

Figure 5 displays the result of applying NonLocal Means Denoising, when applied to the image, this process effectively mitigates the noise that is present in the original image.

Further, data augmentation on HAM-10000 is applied since the CNN learns well if abundant data is provided, images are rotated by 90° clockwise, counterclockwise, 180° and flipped along the x and y -axis. Since the CNN is not affected by the orientation of the image so the dataset is increased by changing the orientation of all lesion images except Melanocytic nevi as it is present in sufficient numbers. The bar graph depicting dataset’s dimension before and after augmentation is shown, respectively in Figs. 6 and 7.

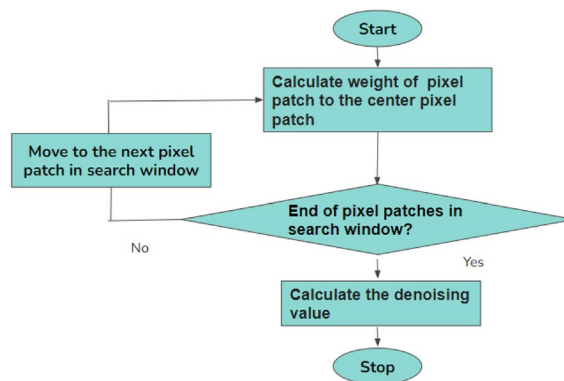


Fig. 4 Flowchart for non-local means denoising

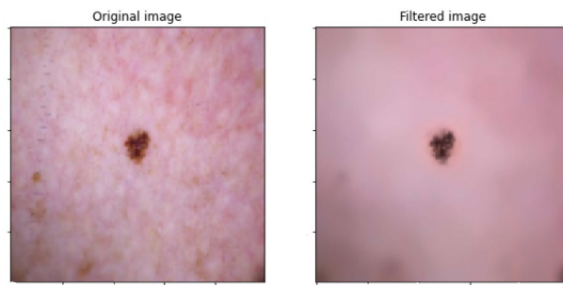


Fig. 5 Filtering using non-local means

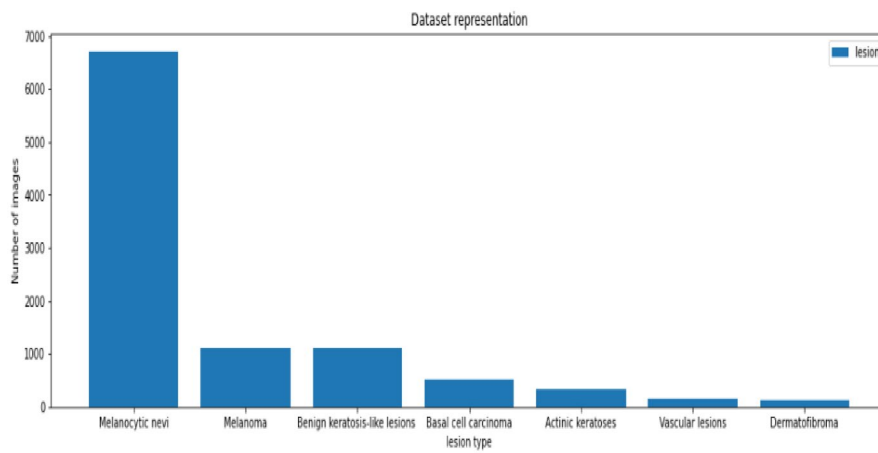


Fig. 6 HAM 10000 before augmentation

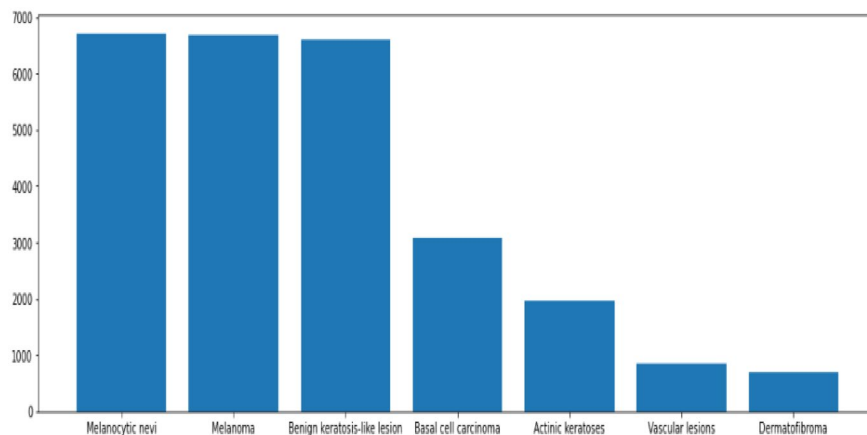


Fig. 7 HAM 10000 after augmentation

Image segmentation and classification

From the dataset, HAM-10000, seven diseases are taken for assessment, and the image count is more than 26,000, after augmentation, which is split into two segments. Here training to testing ratio is maintained 8:2. The ratio was chosen as it is a commonly adopted split in the works of literature, ensuring consistency with previous studies.

Since the HAM-10000 dataset is unbalanced even after augmentation as could be seen in Fig. 7, class weights are assigned within minority classes by allocating higher weights for specific classes and lower weights for other classes.

sci-kit is used to compute the class weights, where an inbuilt function “compute_class_weight” is used. The class weights can be balanced by passing the standard parameter as balanced, which is readily made available in sci-kit-learn models. Following are the class weights provided to the seven classes of diseases in HAM-10000:[0: 0.565, 1: 0.568, 2: 0.575, 3: 1.230, 4: 1.934, 5: 4.454, 6: 5.5].

Further, a second dataset from ISIC-2019 is also included for validation of the model. In this case, the dataset has already been divided into training and testing directories.

In the proposed work, the images are subjected to sparse dictionary learning which is based on the fact that natural signals such as images are allowed for sparse presentation, i.e. expressed as a linear amalgamation of small number of fundamental vectors which are discriminative in nature, and so are used for classification purposes. It plays a crucial role in image denoising and restoration, enabling the recovery of clean images from degraded versions by leveraging sparse representations. Additionally, sparse dictionary learning excels in automatic feature extraction, dimensionality reduction, and robust classification and recognition tasks. It facilitates efficient representation and processing of signals and improves adaptive filtering and compression algorithms by learning tailored dictionaries. By using sparse-based dictionary learning before CNNs, a structured and principled approach to handling complex data representations is ensured, enhancing the performance and accuracy of CNN-based model. This is done using sklearn Mini-batch dictionary learning is a technique for identifying a set of atoms that efficiently sparsely encode the fitted data. The parameter Sparsity controlling parameter, (alpha) is set to one, and the number of dictionary elements to extract are set to 64 with the batch size as 100 for 100 iterations.

The images are then passed to the CNN model for learning and further are tested for the lesion type where its accuracy of prediction is given.

A convolutional layer contains a set of filters that convolve the image with the kernel dot product and compute the image at each spatial location, where the kernel slides by the length stride length. The activation function is the final component of a convolutional layer, which enhances the nonlinearity of the output, which is also referred to as a feature map. Subsequently, the pooling layer is linked to reduce the input size and make the detected features more robust. The batch normalization technique is used to expedite the training process, elevate learning rates, and facilitate learning, it is integrated between the layers of a Neural Network. As a result, rather than utilizing raw data, mini-batches are employed. Furthermore, the optimizers are utilized with the aim of minimizing the losses and to modify various attributes like learning rate, weights of neural networks. The complete flow is explained with visuals in Fig. 8.

Model parameters and implementation

The processing layers encompass Convolution Layer, Pooling Layer, Rectified Linear Units (ReLU), Normalization, Softmax layer, and Fully Connected dense layers.

For the CNN model, the parameters listed below have been utilized.:

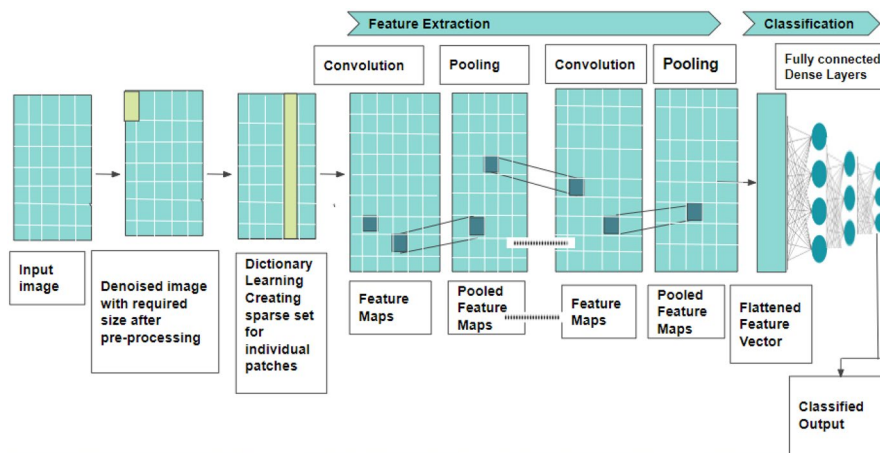


Fig. 8 Overview of the model

- Filters: Here, the suggested approach is to create a bottleneck structure which is to, keep increasing neurons in the first few layers and then reduce it, in our case the order being 128, 256, 512, 512, 256, for better learning.
- Kernel Size: The size of the convolutional filter in pixels. For learning minute details, smaller Kernel Size (1 × 1), (3 × 3) are used, and initially, it has been kept bigger for learning bigger details as (11 × 11) and (5 × 5).
- Activation: The activation functions being used are Relu and Softmax. ReLu is given as-

$$f(x) = \begin{cases} x, & x \geq 0. \\ 0, & x < 0. \end{cases} \tag{7}$$

$$\sigma(z)_i = \frac{e^{z_i}}{\sum_{j \in (1,k)} e^{z_j}} \tag{8}$$

Here in equation (7), function $f(x)$ is defined as x when x is greater than or equal to 0, and 0 when x is less than 0. In equation (8) z is a vector, z_i is the i -th component of z , and k is the number of components in the vector z . ReLU activation function is employed in the hidden layers for avoiding the problem of vanishing gradient and computation of better performance. The Softmax function is employed in the final output layer to transform the network’s output into a probability distribution across the predicted output classes, ensuring normalization..

- Optimizer: Adaptive Moment Estimation (Adam) combines the merits of both momentum-based optimization and root mean square propagation (RMSprop) to achieve fast convergence, low memory requirements, and robustness to noisy gradients, making it a popular choice for deep learning practitioners.
- The determination of batch size is dependent on the number of training data instances that need to be distributed across the neural network during each iteration (train step). An epoch consists of a series of training steps. It takes several epochs to train the neural network so that the detected error approaches zero. So,

choosing an appropriate batch size and epoch count is crucial for achieving optimal performance in neural network training, as smaller batch sizes tend to result in noisier gradients while larger batch sizes can lead to slower convergence and overfitting. The batch size is kept as 32 and epoch as 20 initially for testing denoising and is further epoch set to 100.

After successfully implementing the proposed methodology, the model has demonstrated promising results, confirming its potential to effectively detect skin diseases to a great extent.

Experiments and results

Dataset

HAM-10000, where HAM denotes for “Human Against Machine” [34]. It has total 10,015 images, which is a large consolidation of dermatoscopic pigmented lesion images obtained from different sources. It is an open-source dataset and covers all significant diagnostic categories related to pigmented lesions. The dermatoscopic images have been sourced from diverse populations and acquired and stored using various modalities. The dataset contains seven different classes of skin cancer lesions, i.e. Melanocytic nevi (nv), Melanoma (mel), Benign keratosis-like lesions (bkl), Basal cell carcinoma (bcc), Actinic keratoses (ak), Vascular lesions (vasc), Dermatofibroma (df).

Additionally, this model has also been trained using a dataset provided by The International Skin Imaging Collaboration (ISIC) [35]. ISIC-2019 is freely accessible online for the purpose of research activities, that also contains these seven types of skin lesion images. The dataset is organized into two directories, one for training and another for testing, each of which are further divided into folders containing different lesions as Melanocytic nevi (nv), Dermatofibroma (df), Basal cell carcinoma (bcc), Actinic keratoses (ak), Vascular lesions (vasc), Benign keratosis-like lesions (bkl), Melanoma (mel) containing specific lesion images.

Tools and frameworks

The proposed methodology is initially implemented in Python, using Google Colab.

Further, on the GPU where Nvidia RTX-4000, with 8 GB RAM on Windows 11 is used for final results implementing sparse dictionary learning.

The frameworks used are described as given:

1. Tensorflow: It is a Python library, is an open-source framework created and upheld by Google, designed to efficiently handle matrix data operations and consequently ideal for creating neural networks. It has become the go-to framework for all the tasks involving machine and deep learning.
2. Keras: It is a powerful framework for deep learning that is designed to work with the Python programming language. This library sits “on top” of TensorFlow, offering concise syntax for constructing neural networks, which are then converted into TensorFlow models to facilitate all machine learning operations.
3. Open CV: It is a widely adopted library which is open-source and is utilized for machine learning, computer vision, and image processing. It provides critical sup-

port for real-time operations such as segmentation and detection that are essential in today's systems. When it is integrated with additional libraries like NumPy, Python becomes capable of efficiently analyzing the OpenCV array structure.

4. scikit-learn (sklearn): Sklearn is the one of the best, extremely helpful and reliable Python library for machine learning, offering a myriad of efficient tools employed in machine learning and statistical modeling, involving classification, clustering, regression, and dimensionality reduction.

Experimental setup

In an attempt to measure the performance of the proposed skin disease recognition model, we have analyzed some related research that resembles our research. This work is making a contribution to the machine learning strategy that provides the possibility of detecting a large variety of skin conditions in the human by enlarging the image dataset.

The first step in the technique we implemented is to obtain images of the impacted skin. Usually the images available differ in properties, so preprocessing techniques are performed. The datasets used are standard, HAM- 10000 and ISIC-2019 containing lesion images. Denoising greatly improves the image quality and performance of the model. Following this, sparse dictionary learning further enhances the accuracy of the model. For both datasets, denoising is performed, followed by sparse dictionary learning to improve the fidelity of the model before implementing the CNN for classification.

Experimental analysis

We can see from the resulting plots that without denoising the value accuracy achieved is 56.79% and that with denoising it is 71.30% for 20 epochs in HAM-10000 datasets. Further, the model loss curve has larger spikes in the case of non-denoised images as compared to those denoised using non-local means denoising, this could be seen for both the datasets, as it reduces bad data by fixing poor-quality data with noise, which makes harder for the system to detect the underlying patterns. The initial results without performing denoising are illustrated in Figs. 9 and 11. Further, Figs. 10 and 12 show the plots obtained after denoising. These plots show that the results obtained with denoising are better than those performed without denoising (Figs. 11, 12).

The curves depicting model accuracy and model loss for HAM-10000 are given in Figs. 13 and 14 respectively. The curves depicting model accuracy and model loss for ISIC-2019 datasets are given in Figs. 15 and 16 respectively. The accuracy obtained for HAM-10000 dataset is 85.61% and the accuracy obtained for the ISIC-2019 dataset is 81.23%. This shows that accuracy can be greatly improved using sparse dictionary learning. The relative percentage difference of accuracy for HAM-10000 is 17.27% greater than that obtained with Inception V3 as stated by Eddy et al., and the results for the ISIC dataset are also comparable as could be observed in Table 2.

Other than accuracy other plots obtained are ROC, Confusion Matrix and Heat-Map as given in Figs. 17, 19, 21 and Figs. 18, 20, 22 for HAM-10000 and ISIC dataset respectively. We have obtained ROC Curve and AUC metrics for multiclass classification in both the datasets, as this shows the separability of all the classes by all possible

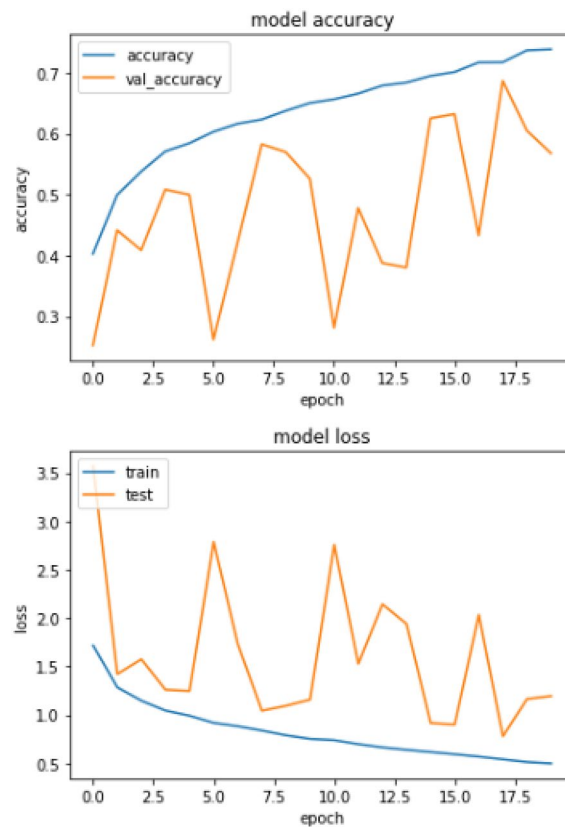


Fig. 9 Without non-local means denoising on HAM-10000

thresholds, i.e. how well the model is classifying each class. The ROC curves have been illustrated in Figs. 17 and 18.

Similarly, the Area Under the Curve (AUC) serves as a metric that evaluates the capability of a binary classifier to differentiate between the classes and serves as an aggregate rendering of the given ROC curve. The model’s effectiveness in discerning between positive and negative classes improves as the AUC rises. The model obtained an AUC score of 0.977 for HAM-10000 and for ISIC-2019 AUC score is 0.963. These results demonstrate the significant impact of denoising and sparse dictionary learning on model performance, with substantial improvements in accuracy and reduced model loss. The high AUC scores and other evaluation metrics further validate the model’s robustness, making it a reliable tool for accurately detecting and classifying various skin conditions.

Another way to understand the effectiveness of categorical classifiers is by using the Confusion Matrix. The confusion matrix obtained for both datasets is given in Figs. 19 and 20. The Heatmap gives a graphical representation of data using color coding to represent the relative values of different data points. The Heatmap obtained for both datasets is given in Figs. 21 and 22 respectively.

While the chosen NLM denoising parameters were effective in reducing noise, these parameters might not be universally optimal across all image variations within the dataset. A fixed patch size and window size may not account for all types of noise, potentially leaving some residual noise unaddressed. Implementing adaptive NLM denoising, where

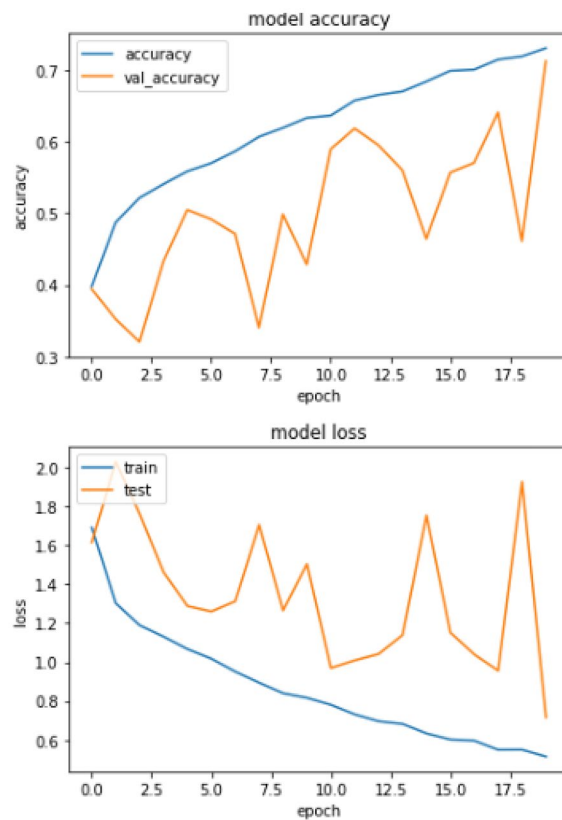


Fig. 10 With non-local means denoising on HAM-10000

the patch size and window size are dynamically adjusted based on the image characteristics, could lead to more effective noise reduction, potentially improving the model’s accuracy. Further hyperparameter tuning within the sparse dictionary learning framework, such as adjusting the number of dictionary atoms or the sparsity level, could optimize the model’s ability to capture underlying patterns, thereby enhancing it’s accuracy. By considering these improvements, the model’s accuracy and robustness could be further enhanced, making it more effective in real-world applications.

Conclusion and future scope

The proposed technique in the manuscript includes the implementation of a computer-based system for the detection of skin condition. To begin with, the images from the database are gathered and subjected to pre-processing by performing denoising of the images, applying non-local means denoising algorithm. The diseases are classified using sparse dictionary learning and the CNN model. The accuracy achieved by the proposed model is 85.61% for HAM-10000 dataset and accuracy attained is 81.23% for the ISIC-2019 dataset.

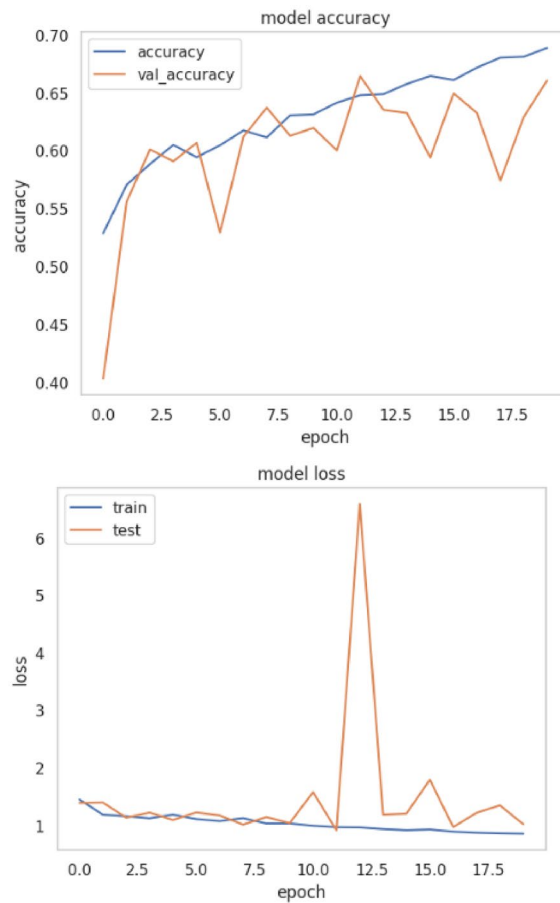


Fig. 11 Without non-local means denoising on ISIC-2019 dataset images

The application of automated diagnostics for skin diseases has the potential to yield significant advantages. However, accurate diagnosis increases the need for a robust automated diagnostic process, that can be used by both the skilled experts and novice clinicians. The skin and skin diseases pattern differs from country to country. This work can be extended by using a diverse dataset with images of all differing skin hues and qualities to further enhance the learning of the model with including more categories of skin diseases to make the model more versatile. The efficiency of the model can be improved with varied hyperparameters of the CNN model for the best possible results. Further, work can be extended by taking real-time images of the lesions and affected skin, so the system could be used in mobile applications and made available to remote people as well.

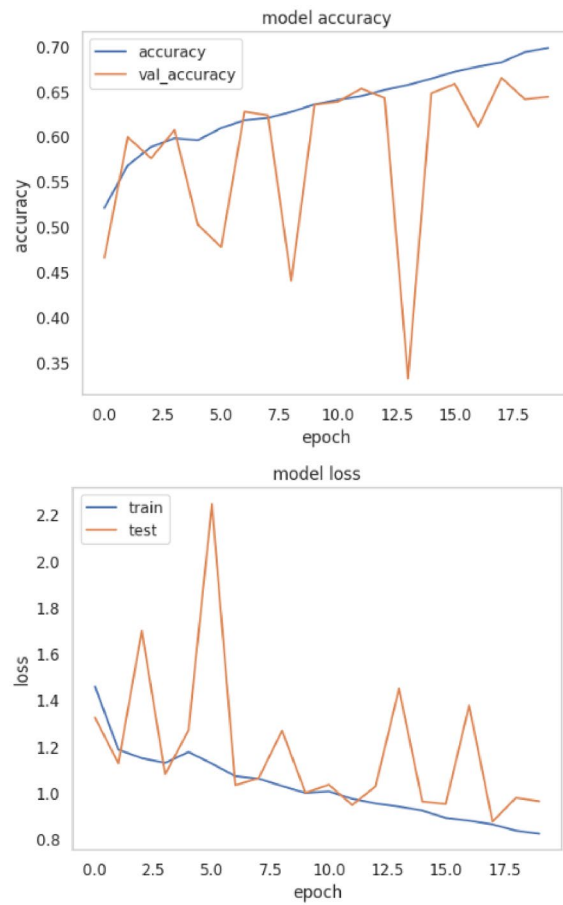


Fig. 12 With non-local means denoising on ISIC-2019 dataset images

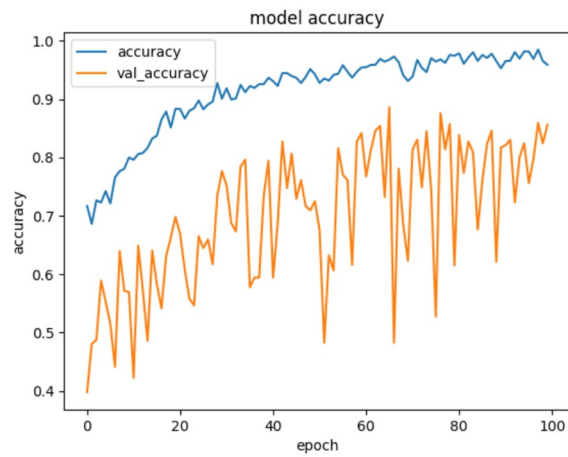


Fig. 13 Model accuracy curve obtained for HAM 10000 with sparse dictionary learning

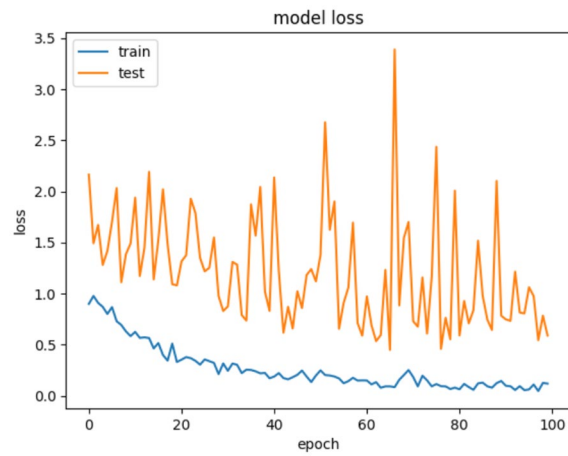


Fig. 14 Model loss curve obtained for HAM 10000 with sparse dictionary learning

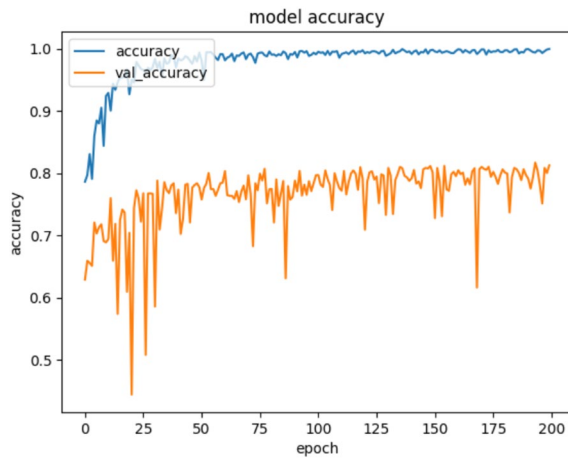


Fig. 15 Model accuracy curve for ISIC-2019 with sparse dictionary learning

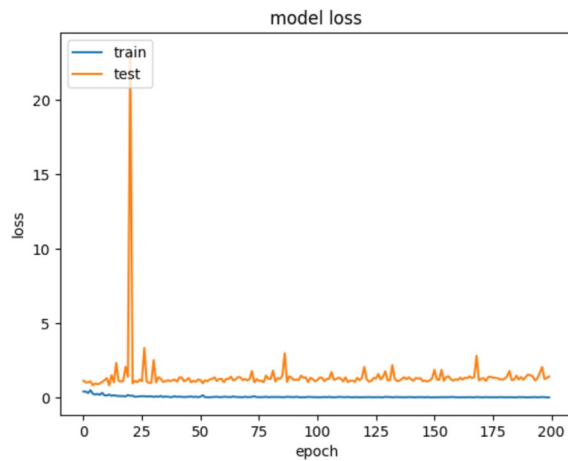


Fig. 16 Model loss curve obtained for ISIC-2019 with sparse dictionary learning

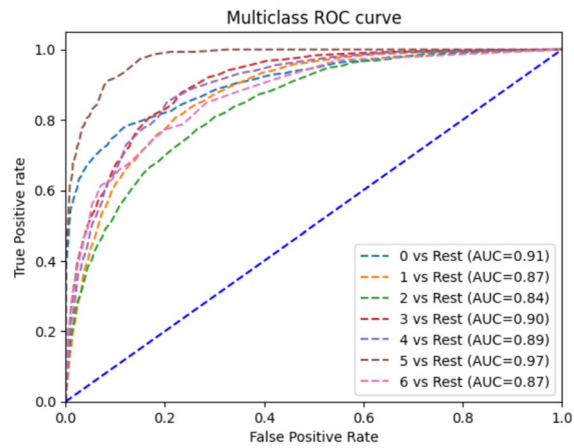


Fig. 17 ROC curve obtained for HAM-10000 where defined classes are 0-nv, 1-mel, 2-bkl, 3-bcc, 4-ak, 5-vasc, 6-df

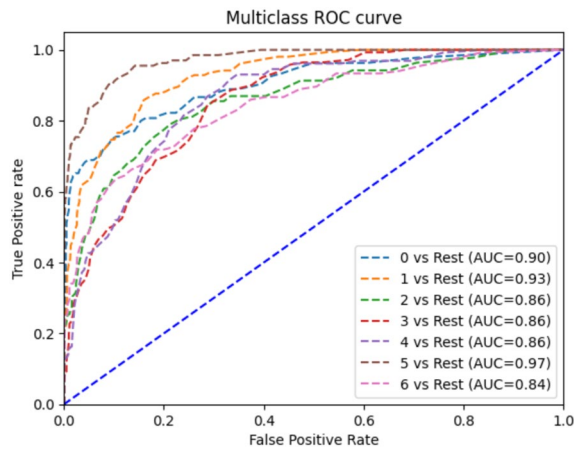


Fig. 18 ROC curve obtained for ISIC-2019 where defined classes are 0-nv, 1-df, 2-bcc, 3-ak, 4-vasc, 5-bkl, 6-mel

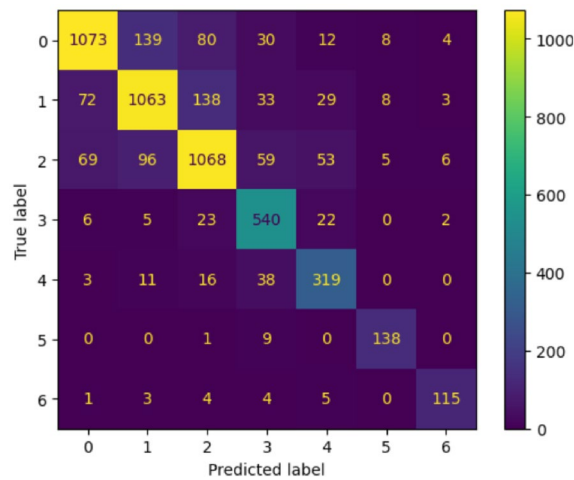


Fig. 19 Confusion Matrix obtained for HAM-10000

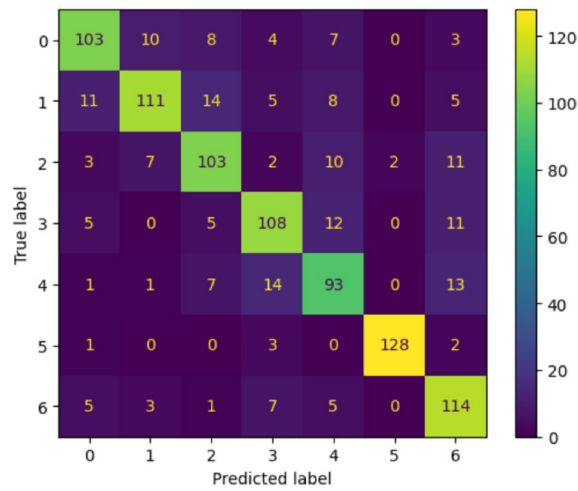


Fig. 20 Confusion Matrix obtained for ISIC-2019

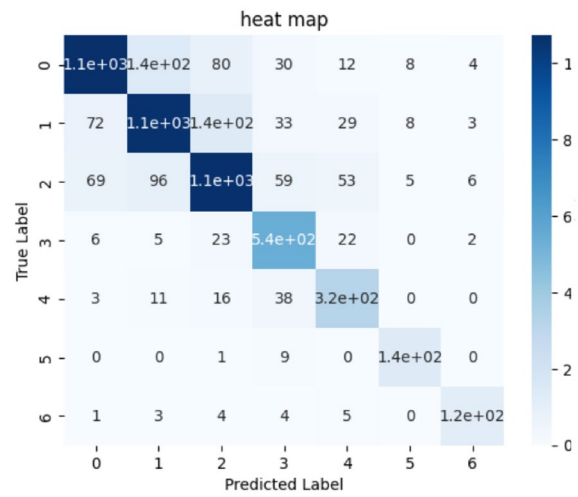


Fig. 21 Heat Map obtained for HAM-10000

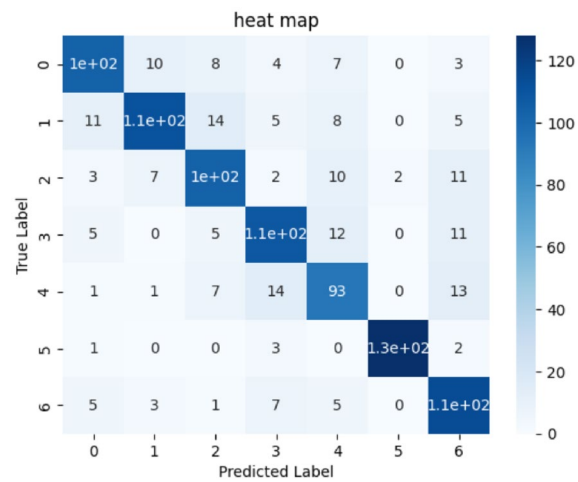


Fig. 22 Heat Map obtained for ISIC-2019

Table 2 Comparison of results with existing State-of-the-Arts

References	Dataset	Techniques used	Results
Eddy et al. [17]	HAM-10000	MobileNet V1 and Inception V3 pretrained CNN models	Inception V3 accuracy-72% MobileNet V1 accuracy-58%
Chaturvedi et al. [18]	HAM-10000	Pre-trained CNN model MobileNet	Overall accuracy-83.1%
Proposed model	HAM-10000	Sparse dictionary based CNN	Accuracy- 85.61%
Thaajwer et al. [8]	ISIC	Shape, color, GLCM with SVM classifier	Overall accuracy-83%
Saranya et al. [20]	ISIC and Demofit	VGG19	Overall accuracy-94.91%
Mijwil et al. [2]	ISIC 2019-2020	(ConvNet)with three architectures InceptionV3, ResNet,VGG19	Best architecture-InceptionV3 accuracy app.-86.90%
Proposed model	ISIC-2019	Sparse dictionary based CNN	Accuracy-81.23%

Author contributions

Apeksha Pandey and Manepalli Sai Teja: Data curation, Software, Writing, Original draft preparation. Parul Sahare: Conceptualization, Methodology, Supervision. Vipin Kamble and Mayur Parate: Visualization, Investigation, Validation, Writing, Reviewing and Editing. Parul Sahare and Mohammad Farukh Hashmi: Writing, Reviewing and Editing. All authors have given approval to the final version of the manuscript.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data availability

Data will be made available on request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

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

Received: 30 January 2024 Accepted: 27 August 2024

Published online: 06 September 2024

References

- Ajith A, Goel V, Vazirani P, Roja MM (2017) Digital dermatology skin disease detection model using image processing. In: International conference on intelligent computing and control systems (ICICCS)
- Mijwil MM (2021) Skin cancer disease images classification using deep learning solutions. *Multimed Tools Appl* 80:26255–26271
- Jain S, Jagtap V, Pise N (2015) Computer aided melanoma skin cancer detection using image processing. *Procedia Comput Sci* 48:735–740
- Garg N, Sharma V, Kaur P (2018) Melanoma skin cancer detection using image processing. *Sensors and image processing*, vol 651. *Advances in Intelligent Systems and Computing*. Springer, Singapore
- Suganya R (2016) An automated computer aided diagnosis of skin lesions detection and classification for dermoscopy images. In: International conference on recent trends in information technology (ICRTIT), pp 1–5
- Dorj U-O, Lee K-K, Choi J-Y, Lee M (2018) The skin cancer classification using deep convolutional neural network. *Multimed Tools Appl* 77:9909–9924
- Hasija Y, Garg N, Sourav S (2017) Automated detection of dermatological disorders through image-processing and machine learning. In: 2017 international conference on intelligent sustainable systems (ICISS), pp 1047–1051
- Thaajwer MA, Ishanka UP (2020) Melanoma skin cancer detection using image processing and machine learning techniques. In: 2020 2nd international conference on advancements in computing (ICAC), pp 363–368
- Nawar A, Sabuz NK, Siddiquee SMT, Rabbani M, Biswas AA, Majumder A (2021) Skin disease recognition: a machine vision based approach. In: 2021 7th international conference on advanced computing and communication systems (ICACCS), pp 1029–1034

10. Thiyaneswaran B, Anguraj K, Kumarganesh S, Thangaraj K (2020) Early detection of melanoma images using gray level co-occurrence matrix features and machine learning techniques for effective clinical diagnosis. *Int J Imaging Syst Technol* 31:682–694
11. Mustafa S, Kimura A (2018) A SVM-based diagnosis of melanoma using only useful image features. In: 2018 international workshop on advanced image technology (IWAIT), pp 1–4
12. Alenezi NSA (2019) A method of skin disease detection using image processing and machine learning. *Procedia Comput Sci* 163:85–92
13. Kumar VB, Kumar SS, Saboo V (2016) Dermatological disease detection using image processing and machine learning. In: 2016 third international conference on artificial intelligence and pattern recognition (AIPR), pp 1–6
14. Yasir MR, Rahman A, Ahmed N (2014) Dermatological disease detection using image processing and artificial neural network. In: 8th international conference on electrical and computer engineering, pp 687–690
15. Ningrum DNA, Yuan S-P, Kung W-M, Wu C-C, Tzeng I-S, Huang C-Y, Yu-Chuan J, Wang Y-C (2021) Deep learning classifier with patient's metadata of dermoscopic images in malignant melanoma detection. *J Multidiscip Healthc* 14:877–885
16. Harangi B (2018) Skin lesion classification with ensembles of deep convolutional neural networks. *J Biomed Inform* 86:25–32
17. Eddy PIK, Kusuma HA, Ratna AAP, Nurtanio I, Hidayati AN, Purnomo MH, Nugroho SMS, Rachmadi RF (2019) Disease classification based on dermoscopic skin images using convolutional neural network in teledermatology system. In: 2019 international conference on computer engineering, network, and intelligent multimedia (CENIM), pp 1–5
18. Chaturvedi SS, Gupta K, Prasad PS (2020) Skin lesion analyser: an efficient seven-way multi-class skin cancer classification using MobileNet. In: Advanced machine learning technologies and applications: proceedings of AMLTA 2020
19. Saranya K, Vijayashaarathi S, Sasirekha N, Rishika M, Rajeswari PSR (2024) Skin disease detection using CNN (convolutional neural network). In: International conference on data engineering and communication systems (ICDECS)
20. Gupta M, Kumar R, Nandan Pradhan AK, Obaid AJ (2024) Skin disease detection using neural networks. In: International conference on advancements in smart, secure and intelligent computing (ASSIC)
21. Ahalya RK, Babu G, Sathish S, Shruthi K (2024) Automated skin disease detection using deep learning algorithms. In: International conference on communication and signal processing (ICCCSP)
22. Jain A, Rao ACS, Jain PK, Abraham A (2022) Multi-type skin diseases classification using OP-DNN based feature extraction approach. *Multimed Tools Appl* 81:6451–6476
23. Nawaz M, Mehmood Z, Nazir T, Naqvi RA, Rehman A, Iqbal M, Saba T (2022) Skin cancer detection from dermoscopic images using deep learning and fuzzy k-means clustering. *Microsc Res Tech* 85(1):339–351
24. Yap J, Yolland W, Tschandl P (2018) Multimodal skin lesion classification using deep learning. *Exp Dermatol* 27(11):1261–1267
25. Ge Z, Demyanov S, Chakravorty R, Bowling A, Garnavi R (2017) Skin disease recognition using deep saliency features and multimodal learning of dermoscopy and clinical images. In: Medical image computing and computer assisted intervention MICCAI
26. Yilmaz E, Trocan M (2021) A modified version of GoogLeNet for melanoma diagnosis. *J Inf Telecommun* 5:395–405
27. Chaturvedi SS, Tembhurne JV, Diwan T (2020) A multi-class skin cancer classification using deep convolutional neural networks. *Multimed Tools Appl* 79:28477–28498
28. Karunanayake RK, Dananjaya WM, Peiris MY, Gunatileka BR, Lokuliyana S, Kuruppu A (2020) CURETO: skin diseases detection using image processing and CNN. In: 2020 14th international conference on innovations in information technology (IIT), pp 1–6
29. Rimi TA, Sultana N, Foysal MFA (2020) Derm-NN: skin diseases detection using convolutional neural network. In: 2020 4th international conference on intelligent computing and control systems (ICICCS), pp 1205–1209
30. Pangti R, Mathur J, Chouhan V, Kumar S, Rajput L, Shah S, Gupta A, Dixit A, Dholakia D, Gupta S, Gupta S, George M, Sharma VK, Gupta S (2020) A machine learning-based, decision support, mobile phone application for diagnosis of common dermatological diseases. *Acad Dermatol Venereol* 35(2):536–545
31. Diwan T, Shukla R, Ghuse E, Tembhurne JV (2023) Model hybridization & learning rate annealing for skin cancer detection. *Multimed Tools Appl* 82:2369–2392
32. Sharma AK, Tiwari S, Aggarwal G, Goenka N, Chakrabarti AKP, Chakrabarti T, Gono R, Leonowicz Z, Jasinski M (2022) Dermatologist-level classification of skin cancer using cascaded ensembling of convolutional neural network and handcrafted features based deep neural network. *IEEE Access* 10:17920–17932
33. Li H, Pan Y, Zhao J, Zhang L (2021) Skin disease diagnosis with deep learning: a review. *Neurocomputing* 464:364–393
34. Tschandl P, Rosendahl C, Kittler H (2018) The HAM10000 dataset, a large collection of multi-source dermoscopic images of common pigmented skin lesions. *Sci Data* 5:1–9
35. Combalia M, Codella NC, Rotemberg V, Helba B, Vilaplana V, Reiter O, Carrera C, Barreiro A, Halpern AC, Puig S, Malvehy J (2019) Isic-2019 challenge. *Bcn20000: dermoscopic lesions in the wild*

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.