

A machine learning approach for skin disease detection and classification using image segmentation

Mostafiz Ahammed ^a, Md. Al Mamun ^{b,*}, Mohammad Shorif Uddin ^a

^a Department of Computer Science and Engineering, Jahangirnagar University, Dhaka, Bangladesh

^b Department of Public Health and Informatics, Jahangirnagar University, Dhaka, Bangladesh



ARTICLE INFO

Keywords:

Skin Disease
Machine Learning
Image Segmentation
Decision Tree
Support Vector Machine
K Nearest Neighbor (KNN)

ABSTRACT

Skin diseases are common health problems around the world. The perils of the infections are invisible, which cause physical health distress as well as initiate mental depression. In addition, it sometimes leads to skin cancer in severe cases. Subsequently, diagnosing skin diseases from clinical images is one of the foremost challenging tasks in medical image analysis. Moreover, when performed manually by medical experts, diagnosing skin diseases is time-intensive and subjective. As a result, both patients and dermatologists require automatic skin disease prediction, which makes the treatments plan faster. In this work, we introduce a digital hair removal technique based on morphological filtering such as Black-Hat transformation and inpainting algorithm and then apply Gaussian filtering to de-blur or denoise the images. In addition, we apply the automatic Grabcut segmentation technique to segment out the affected lesions. For extracting underlying input patterns from the skin images, we apply the Gray Level Co-occurrence Matrix (GLCM) and statistical features techniques. Three computationally efficient machine learning techniques, Decision Tree (DT), Support Vector Machine (SVM), and K-Nearest Neighbor (KNN) classifiers are applied using the extracted features for effectively classifying the skin images as melanoma (MEL), melanocytic nevus (NV), basal cell carcinoma (BCC), actinic keratosis (AK), benign keratosis (BKL), dermatofibroma (DF), vascular lesion (VASC), and Squamous cell carcinoma (SCC). The models are validated using two standard datasets ISIC 2019 challenge and HAM10000. SVM performs slightly better than the other two classifiers. We have also compared our work with state-of-the-art methods.

1. Introduction

The human body is made up of several organs. Skin is one of them. It is the largest organ covering the entire human body [1]. Any disorder that affects human skin is called skin disease [2]. Skin disease is one of the most contagious diseases in the world. According to the World Health Organization (WHO), about 2794 persons in Bangladesh died from skin cancer in 2018 [3]. WHO also said that more than 14 million cases were diagnosed, and 9.6 million deaths occurred globally in 2018 [4]. It is the change of color or texture of the skin. The causes of skin diseases are viruses, bacteria, allergy, or fungal infections [5]. The genetic factor also causes skin disorders. Generally, skin disease occurs in the thin outer layer of the skin, called epidermis can be visualized by human eyes that cause psychological depression and lead to physical injuries.

There are different types of skin lesions: Actinic keratosis (AK), Basal cell carcinoma (BCC), Benign keratosis (BKL), Dermatofibroma (DF), Melanoma (MEL), Melanocytic nevus (NV), Squamous cell carcinoma (SCC), and Vascular lesion (VASC), are shown in Fig. 1. The

lesions are different in terms of their symptoms and severity. Some are permanent, and some are temporary and may be painless or painful. Among these skin diseases, melanoma is the most deadly and dangerous type. However, about 95% of skin disease patients can be recovered if identified at an initial state. An automatic computer-aided system can be beneficial to classify skin diseases accurately.

There is a massive gap between dermatologists and skin disease patients as many people do not know the types, symptoms, and stages of skin disease. Sometimes it requires a long time to show the signs. For this, it requires early and quick detection. But it may be difficult and expensive to diagnose skin diseases correctly to identify the type and stage of the disease. The automatic computer-aided system based on machine learning approaches has made it possible to detect the types of skin disease more accurately and quickly.

Many researchers have worked on skin disease classification for the last three decades. The area is so significant and has become a hot research topic. Even though many research papers are done on skin disease detection and classification, there is still a gap to be filled. Most of the previous work is based on a single disease [6,7], and those that

* Corresponding author.

E-mail addresses: mostafizs154@gmail.com (M. Ahammed), almamun@juniv.edu (M.A. Mamun), shorifuddin@juniv.edu (M.S. Uddin).

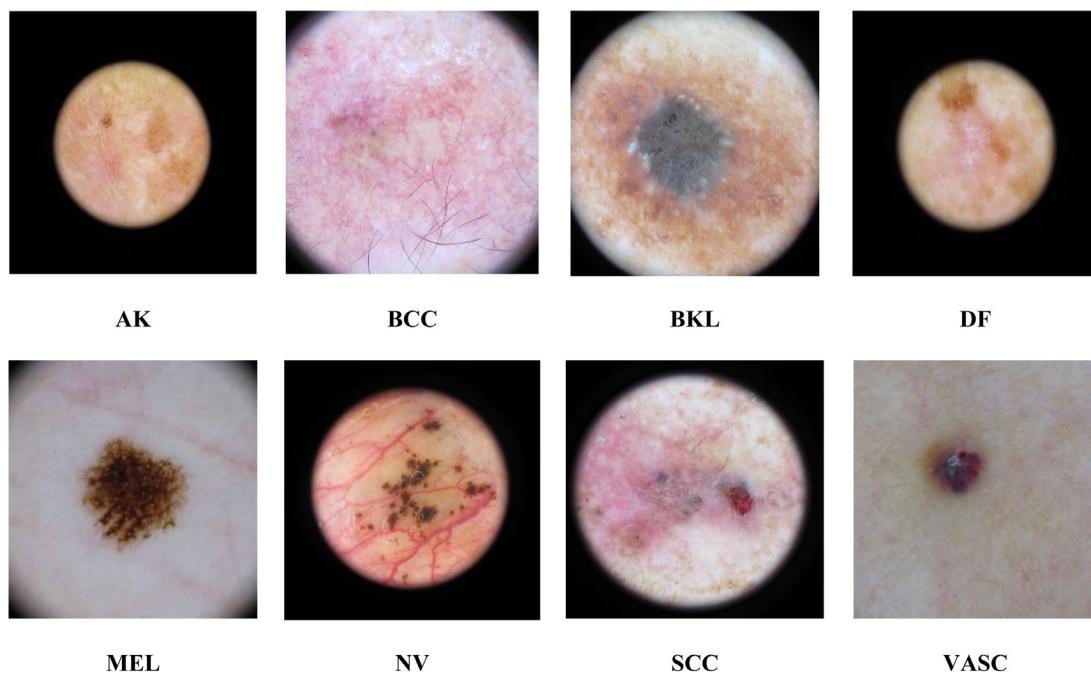


Fig. 1. Sample of Skin Disease Images collected from the ISIC 2019 Challenge Dataset.

have been done are inadequate for classifying multiple classes [8]. The classification task of multiple classes is very challenging as the skin disease presents more similar behavior.

The main contribution of this research is outlined as follows:

- To build a model for removing hair using Black-Hat Transformation and Image Inpainting algorithm.
- To develop a powerful segmentation model using the Grabcut technique that detects the lesion without losing any information and makes the images more suitable for further processing.
- To develop an automatic classification model for skin diseases classification based on a sufficient number of relevant features with high accuracy.

The remaining of the research is structured as follows: part 2 discusses the related literature and the problem statement. The proposed methodology is described in part 3, including preprocessing, skin lesion segmentation, feature extraction, and skin disease classification. The datasets used in this work and data preparation tasks are discussed in part 4. The experimental results of the proposed approach are discussed in part 5. Quantitative analysis and discussion on performance evaluation are also included in this chapter. Finally, part 6 concludes the study with future prospects.

2. Literature Review

Many researchers proposed several approaches for the classification of skin diseases. Related works are categorized into different types based on datasets, feature extraction techniques, feature selection techniques, and classification models. This section examined several relevant research articles to discover previous studies' tools and procedures and identify research gaps.

Jagdish et al. [9] proposed a model for skin disease detection using image processing methods. They applied fuzzy clustering on 50 sample images with KNN and SVM classification algorithm with wavelet analysis. They showed that the K-Nearest Neighbor classification algorithm works well compared to the Support vector machine (SVM) classification technique with an accuracy of 91.2% and the algorithm identified the type of skin disease using classification methods. But they

worked with only 50 sample images containing two classes (basal and squamous disease).

Naeem et al. [10] proposed a model to predict skin cancer using image processing strategies and support vector machines (SVMs). They used various pre-handling procedures for clamor evacuation and picture improvement and GLCM method to separate a few highlights in the image. Finally, the classifier classified the images as harmful or harmless.

Bandyopadhyay et al. [11] proposed a model combining deep learning (DL) and machine learning (ML). They applied deep neural networks Alexnet, Googlenet, Resnet50, and VGG16 for feature selection and Support Vector Machine, Decision tree, and Ensemble boosting Adaboost classifier for classification. Finally, they carried out a comparative study to identify the best prediction model.

Kalaivani et al. [12] proposed a new method that combines two separate data mining approaches into a single unit, as well as an ensemble approach that combines both data mining techniques into a single group. They applied the ensemble deep learning technique on an informative Dermatology publicly accessible dataset ISIC2019 image and categorized skin disorders into seven categories. They observed that the ensemble technique predicted skin diseases more accurately and effectively.

AlDera et al. [13] presented a skin disease diagnosis model that take an affected skin image and diagnose acne, cherry angioma, melanoma, and psoriasis. They applied Otsu's method for image segmentation and Gabor, Entropy and Sobel techniques for feature extraction on the dermnet NZ and atlas dermatologico. Finally, they applied Support Vector Machine (SVM), Random Forest (RF), and K-Nearest Neighbor (K-NN) classifiers for classification, and achieved 90.7%, 84.2%, and 67.1% accuracy, respectively.

Kshirsagar et al. [14] proposed a system for skin disease classification using MobileNetV2 and LSTM. They primarily aimed on the accuracy in skin disease forecasting for the system and ensured excellent efficiency in storing complete state information for exact forecasts. They also compared with several other traditional models such as CNN and FTNN and observed that the proposed model outperformed others in classifying skin disease and evaluating the progression of tumor development using texture-based information.

Hatem [15] proposed a system for identifying skin lesions and classifying them as normal and benign. He applied k nearest neighbor

(KNN) classifier for classifying skin lesions and achieved an accuracy of 98% only for the two classes.

Kethana et al. [16] proposed a model utilizing Convolutional Neural Network (CNN) for the classification of skin disease. They used a dataset of 10015 images collected from ISIC 2019 dataset and achieved an accuracy of 92% in classifying skin diseases as Melanoma, Nevus, and Seborrheic Keratosis.

Yao et al. [17] proposed a novel single-model based approach for classifying skin lesions on small and imbalanced datasets. They first trained various Deep Convolutional Neural Network (DCNN) on various small and imbalanced datasets and then added regularization DropOut and DropBlock to reduce overfitting and proposed a Modified RandAugment augmentation strategy to deal with the defects of sample underrepresentation in the small dataset. Finally, they applied a novel Multi-Weighted New Loss (MWNL) function and an end-to-end cumulative learning strategy (CLS) to lessen the effect of abnormal samples on training and to deal with the problem of unequal sample size and classification difficulty. This model achieved a high classification performance at a low cost of computational resources and inference time.

Padmavathi et al. [18] proposed an automated skin lesion classification approach utilizing a pre-trained and fine-tuned deep learning network. They assessed the performance using well-known quantitative criteria such as specificity, sensitivity, precision, and accuracy and compared using various transfer learning methods.

Maduranga et al. [19] presented an artificial intelligence (AI) based mobile application for the detection of skin diseases. They applied the Convolutional Neural Networks (CNN) on HAM10000 dataset for classifying skin disease. They also built a mobile application for quick and accurate identification using MobileNet with transfer learning and obtained an accuracy of 85%.

Jain et al. [20] proposed a novel optimal probability-based deep neural network (OP-DNN) to appropriately diagnose the skin disease. Initially, they removed the unwanted contents from the images and then extracted features for training the Optimal Probability-Based Deep Neural Networks (OP-DNN). They also applied an optimization approach and obtained an accuracy of 95%, 0.97 of specificity and 0.91 of sensitivity.

Soliman [5] applied a multiclass Support Vector Machine (SVM) for classifying skin diseases where a pre-trained convolutional neural network (CNN) was used for extracting the relevant features. As a result, the model obtained an average accuracy of 100% for classifying only the three different diseases like melanoma, eczema, and psoriasis. But the study deals with only three diseases. So, it is not always the generalized model for various skin diseases. It is the shortcoming of the study.

Janney et al. [21] applied the Support Vector Machine (SVM), the Artificial Neural Network (ANN), and the Naïve Bayes Classifier. The features were extracted from the dermoscopic images using the Gray Level Co-occurrence Matrix (GLCM), texture, and wavelet features. Finally, the performance of the classification models was compared based on the accuracy, precision, and recall values. The system achieved an average accuracy of 89%, 71%, and 71% for ANN, SVM, and Naïve Bayes classifiers, respectively. But, the system has a poor performance on SVM and Naïve Bayes Classifiers. It is a vital drawback of the study.

Alquran et al. [22] used the Support Vector Machine (SVM) classifier for classifying skin images into two categories: normal and abnormal. Then, the Gray Level Co-occurrence Matrix (GLCM) and Asymmetry, Border, Color, Diameter (ABCD) were used for extracting the features from the skin images, and the most appropriate features were selected applying the Principal Component Analysis (PCA) technique. As a result, the system achieved an average accuracy of 92.1%. But the drawback of the system is that it is suitable for only two classes.

Balaji et al. [23] proposed a skin disease detection and segmentation model based on the dynamic graph cut algorithm to segment the lesion of the images and the Naïve Bayes classifier for classifying the types

of diseases. The model outperformed some state-of-the-art approaches, including FCN by 6.5% and SegNet by 8.7% for classifying the diseases on the ISIC 2017 dataset.

Sinthura et al. [24] proposed an advanced skin disease diagnosis model utilizing image processing. First, they applied Otsu's method to segment the disease portion. Next, some GLCM features such as area, perimeter, mean, and entropy were extracted. Finally, they applied the SVM classifier to identify the four classes of diseases: Acne, Psoriasis, Melanoma, and Rosacea. The method identifies skin diseases with an accuracy of 89%. They applied the model to a small dataset containing 100 images and used only a small number of features. These are the drawbacks of the system.

Kumar et al. [25] presented a method for determining whether or not a particular sample is impacted by melanoma. This study's steps are as follows: gathering labeled data from preprocessed images, flattening those images and getting the pixel intensities of images into an array, appending all such arrays into a database, training the SVM with labeled data using a suitable kernel, and using the trained data to classify the samples successfully. The classification accuracy is around 90%. Unfortunately, they only identified a single disease. It is the main drawback of the system.

Hameed et al. [26] proposed an intelligent diagnosis scheme for multiclass skin lesion classification. The proposed methodology is developed using a hybrid approach, including a deep convolutional neural network and an error-correcting output code (ECOC) support vector machine (SVM). The proposed method is intended to categorize skin lesion images into five categories: healthy, acne, eczema, benign, or malignant melanoma. Experiments were carried out on 9,144 images gathered from various sources. The features were extracted using AlexNet, a pre-trained CNN model. The ECOC SVM classifier was utilized for classification. The overall accuracy obtained with ECOC SVM is 86.21%.

Shanthi et al. [27] proposed a method to detect four types of skin diseases using computer vision. The proposed approach involves Convolutional Neural Networks with around 11 layers viz., Convolution Layer, Activation Layer, Pooling Layer, Fully Connected Layer, and Soft-Max Classifier. The architecture is validated using images from the DermNet database. The database contains all forms of skin disorders, but they have focused on four major types of skin diseases: acne, keratosis, eczema herpeticum, and urticaria, each having 30 to 60 different samples. The accuracy of the proposed CNN Classifier ranges from 98.6% to 99.04%. But they worked with only a few images having four classes. These are the main shortcomings of this study.

Bhavani et al. [28] presented a method for detecting various types of dermatological skin disorders using computer vision-based approaches. Three deep learning algorithms, Inception v3, Mobilenet, and Resnet, are utilized for feature extraction in a medical image. A machine learning technique, Logistic Regression, trains and evaluates the medical images. Significant efficiency may be realized by combining the architectures of the three convolutional neural networks. However, the model used in this study was computationally complex, and the dataset contained only three types of skin diseases; the system may not be efficient for multiclass classification.

Hameed et al. [29] presented a novel Computer-aided Diagnosis (CAD) system for diagnosing the most common skin lesions (acne, eczema, psoriasis, benign and malignant melanoma). The applied Dull Razor method, Gaussian filtering for preprocessing, and Otsu's thresholding method for extracting the region of interest from which the features are extracted. Several colors and texture features were extracted and applied to the Support Vector Machine (SVM) with the quadratic kernel. The experiment was performed on 1800 images, and 83% accuracy was achieved for six-class classification. However, the dataset was small, which is a limitation of the system.

Hameed et al. [30] presented a method that was an extension of an existing work where a novel classification scheme is proposed for multiclass classification. The proposed classification model includes

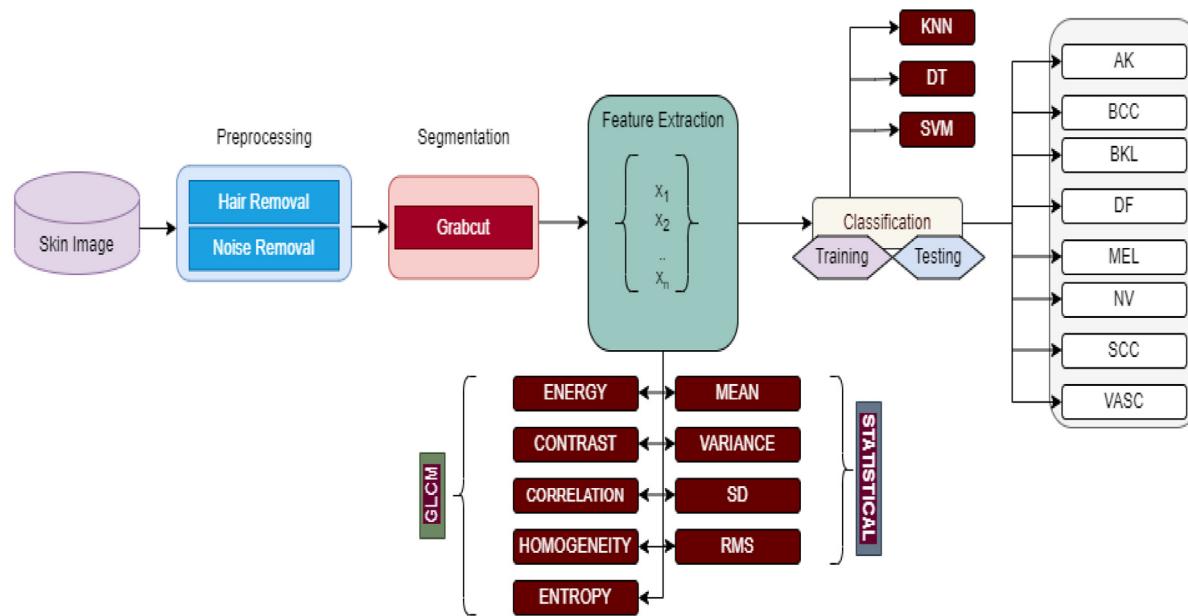


Fig. 2. A Generalized Block Diagram of the Skin Diseases Classification System.

four steps, i.e., preprocessing, segmentation, feature extraction, and classification. About 1800 images of six classes were used in the experiments. After utilizing 10-fold cross-validation, the Quadratic SVM achieved an average accuracy of 94.74%.

Ubale et al. [31] proposed a color phase model for detecting and classifying different skin diseases. They applied the HSV color phase model and LAB color phase model to extract features from the images. Finally, they employed the K Nearest Neighbor (KNN) classification method to classify skin diseases (Acne, Papillomas, Psoriasis, Melanoma, Mycosis, Vitiligo, Warts) that achieved 91.80% accuracy for HSV Color Phase Model and 81.60% accuracy for LAB Color Phase Model. Unfortunately, the preprocessing is not good enough, and there is no verified dataset used in this study.

Albawi et al. [32] proposed unique skin disease identification for three types of skin diseases: Melanoma, Nevus, and Atypical. In the preprocessing step, they applied an adaptive filtering method to remove unwanted noisy areas from the input skin image. Next, an adaptive region growing technique was employed for efficient localization and region of interest (ROI) extraction of disease areas. They used a hybrid feature extraction method composed of two-dimension discrete wavelet transform (2D-DWT) and geometric and texture features to extract relevant features. Finally, convolutional neural networks (CNN) were applied to the International Skin Imaging Collaboration (ISIC) dataset. The proposed method obtained an accuracy of 96.768% for classifying the diseases.

Ozkan et al. [33] presented a model for classifying skin lesions in three groups: normal, abnormal, and melanoma. They applied four machine learning methods: ANN, SVM, KNN, and DT to the PH² dataset. The model achieved an accuracy of 92.50%, 89.50%, 82.00%, and 90.00% for ANN, SVM, KNN, and DT, respectively. However, the dataset contains only three classes; the system may not be efficient for multiclass classification.

From the above discussions, it can be said that most of these existing studies have been performed to predict depression among a specific group of people, like: among people of a particular range of age, patients diagnosed with a particular disease, etc. This study tries to overcome this restriction by considering different socio-demographic information of the people of diverse age ranges, health conditions, and socio-economic statuses.

From the above discussions, it can be concluded that most existing research has been conducted to detect skin diseases of a few classes. Most of the works are based on a single disease. Only a few that have been done are inadequate for classifying multiple classes. Most of them are performed on small size datasets. The preprocessing techniques are not good enough, and thresholding segmentation, edge-based segmentation, clustering-based segmentation are used in most of the studies. This study attempts to solve this limitation by examining the International Skin Imaging Collaboration (ISIC-2019) images, which consists of eight skin diseases, and applying a digital hair removal technique for preprocessing and graph-based segmentation.

3. Proposed Methodology

This chapter described and discussed the proposed approaches for classifying skin diseases. The whole process comprises the following parts: The first step is image preprocessing, the second step is image segmentation, the third step is feature extraction, and the final step is classification. Fig. 2 depicts a high-level overview of the proposed strategy.

3.1. Image Preprocessing

Image preprocessing is converting an image into a more suitable and usable form. The skin images may contain unwanted hair, noise, or distortion. The performance of an image processing system depends on the quality of images. The skin images need to be processed to achieve better performance for the skin disease detection system. It improves the image quality for further processing, reduces complexity, and increases accuracy. The preprocessing steps include the following sections: Image Resizing, Hair Removal, and Noise Removal.

3.1.1. Image Resizing

The images of different sizes do not always get the same number of features. Therefore, the input images increase or decrease in size to resolve this problem. It also shortens the processing time and improves the system's overall performance. In this work, we resized all the input images into a size of 512 × 512. The original image (before resizing) and the scaled image is shown in Figs. 3(a) and 3(b), respectively.

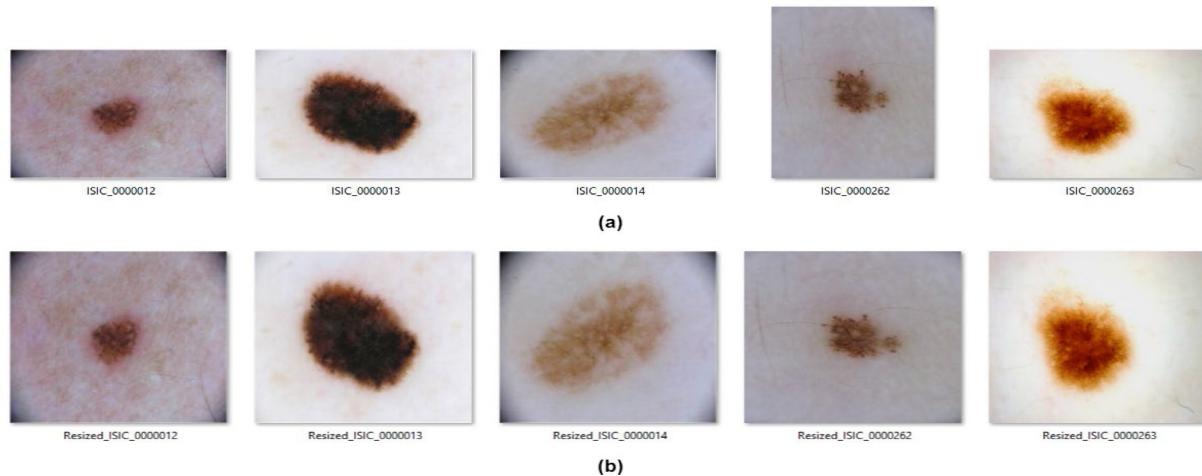


Fig. 3. The figure shows image preprocessing (a) Before Resizing (b) After Resizing.

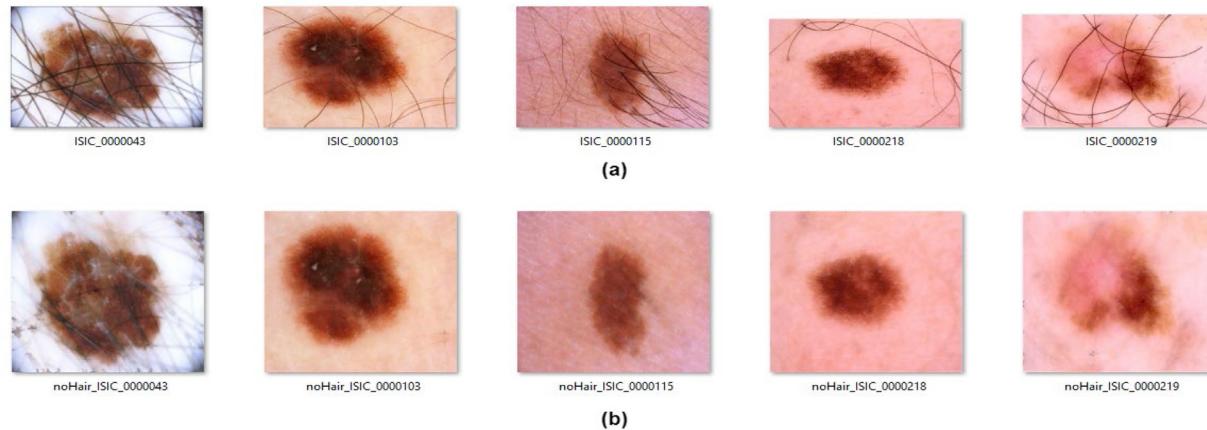


Fig. 4. Preprocessing of Images (a) Before Removing Hair (b) After Removing Hair.

3.1.2. Hair Removal

By removing hair from the image, skin disease detection can be improved. Several techniques such as median filter [34], adaptive threshold [35], Gabor filtering and PDE-based inpainting [36], and morphological operations such as Top Hat filter [37,38] were widely used for removing hairs from skin images. We have used a digital hair removal (DHR) algorithm based on morphological filterings such as Black-Hat transformation and inpainting algorithm to remove hairs in our proposed system. The steps of this DHR algorithm are given below:

1. Convert RGB images into grayscale images
2. Apply Morphological Black-Hat transformation on the grayscale images
3. Create a mask for the inpainting task
4. Apply inpainting algorithm on the original image using this mask

The image before and after removing the hair is shown in Figs. 4(a) and 4(b), respectively.

3.1.3. Noise Removal

Noise can appear in digital images during image acquisition, transmission, and other processing processes. In addition, the quality of images is degraded by an unexpected change of brightness or color information in images. Many filtering techniques such as mean blurring, median blurring, Gaussian blurring, and bilateral filtering were commonly utilized for blurring images or removing noise from images [39]. In this system, we have used Gaussian filtering for blurring images and

removing noise. It applies a kernel for convolutional operations to the images. The kernel values have a Gaussian distribution which is given in Eq. (1).

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (1)$$

Where μ is the mean value and σ is the standard deviation of the distribution function.

We used a 7×7 kernel, where the sigma value is calculated from the kernel. A sample image from the dataset before and after Gaussian filtering is shown in Figs. 5(a) and 5(b).

3.2. Segmentation

The process of segmenting an image into non-overlapping groups or regions is known as image segmentation. It is based on gray level, brightness, color, contrast, texture, and other properties. It separates the identical lesions from the healthy skin around them. It is the most crucial step in effectively evaluating images because it influences the accuracy of the subsequent processes. However, accurate segmentation in microscopic images is challenging due to the vast differences in size, shape, and color of the lesions. In addition, there is a lack of contrast between the lesions and the healthy skin around them. Several segmentation techniques are classified into threshold-based, region-based, and cluster-based, and edge-based image segmentation [40–43].

In this work, we have applied the automatic Grabcut technique for image segmentation based on the k-means clustering and the Hue Saturation Value (HSV) color space. We have used the Grabcut algorithm

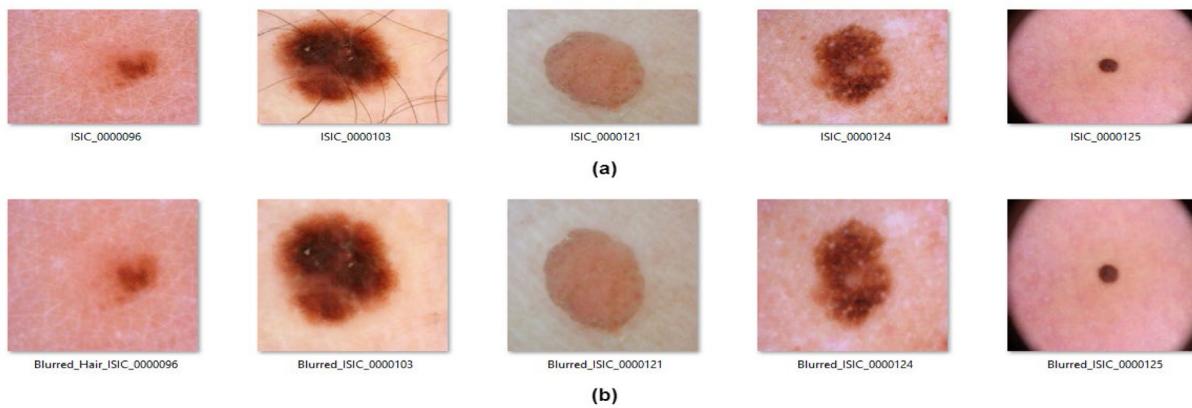


Fig. 5. Preprocessing of Images (a) Before Gaussian Filtering (b) After Gaussian Filtering.

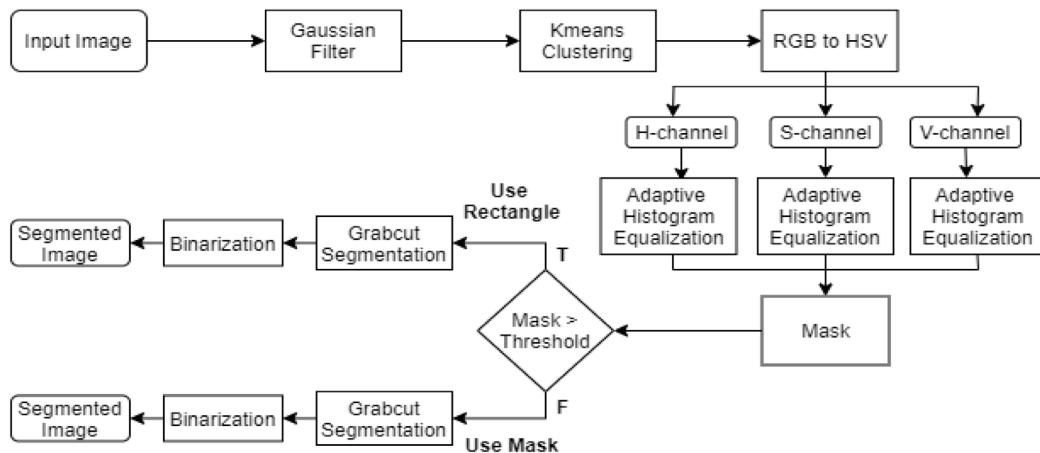


Fig. 6. Block diagram of automatic Grabcut segmentation.

as it estimates the color distribution of the target object and the background using a Gaussian Mixture Model (GMM), which automatically creates a rectangle and extracts the foreground from the background images minimizing user interaction.

3.2.1. Proposed Automatic Grabcut Segmentation

Our proposed automatic Grabcut algorithm performed Grabcut by combining two approaches and using a single threshold parameter. Thus, the suspected lesion region was first retrieved as a mask and then used for Grabcut segmentation. However, if the extracted mask is greater than the threshold, the extracted mask is more than enough, and a rectangle is created to perform the Grabcut. The working of the algorithm is shown using a block diagram given in Fig. 6.

The majority of the skin lesions appear green in HSV color space. Therefore, the green color was extracted from the image and applied to obtain a thresholding mask. When the mask has a value of 1, the image graph considers it to be the foreground, and where it has a value of 0; the image graph considers it the background. Then, we performed the Grabcut segmentation using this mask. However, in other circumstances, the idea of extracting a mask based on the green channel failed due to a tiny difference in intensity between the skin and the lesion, especially when the lesion is large or small.

One thresholding technique was proposed to overcome this problem. The mask dimension and intensity value were used to compute the threshold. The threshold was determined because the mask might contain up to 70% of the total pixels in the image. The mask was rejected if it considered more than 70% of the pixels in the image to be lesion pixels. The visual observation of the dataset assumes an

estimation of these 70% pixels. A small amount of lesion occupies more than 70% of the pixels in the image.

A rectangle was generated if the mask exceeded the threshold, and the rectangle was utilized for Grabcut segmentation. The rectangle dimensions are shown in Eqs. (2) and (3).

$$H_r = H_i - (0.3 \times H_i) \quad (2)$$

$$W_r = W_i - (0.3 \times W_i) \quad (3)$$

Where H_r is the rectangle height, H_i is the image height, W_r is the rectangle width, and W_i is the image width.

Everything outside the rectangle was considered background, and Grabcut segmentation was performed.

A sample image from the dataset before and after segmentation is shown in Figs. 7(a) and 7(b).

3.3. Feature Extraction

Feature extraction is essential for studying and discovering the underlying relationships between various objects. The image categorization, prediction, and recommendation algorithms cannot comprehend the images directly. As a result, feature extraction is necessary to convert them into usable forms. The dermoscopic image has various characteristics that are utilized to describe the image. However, not all characteristics apply to the categorization of skin disease. As a result, the classifier becomes complex and takes more computational effort in several irrelevant features, reducing the classification accuracy. In skin cancer pictures, the best features must reflect the characteristics of the areas. As a result, a sufficient number of features should be retrieved to

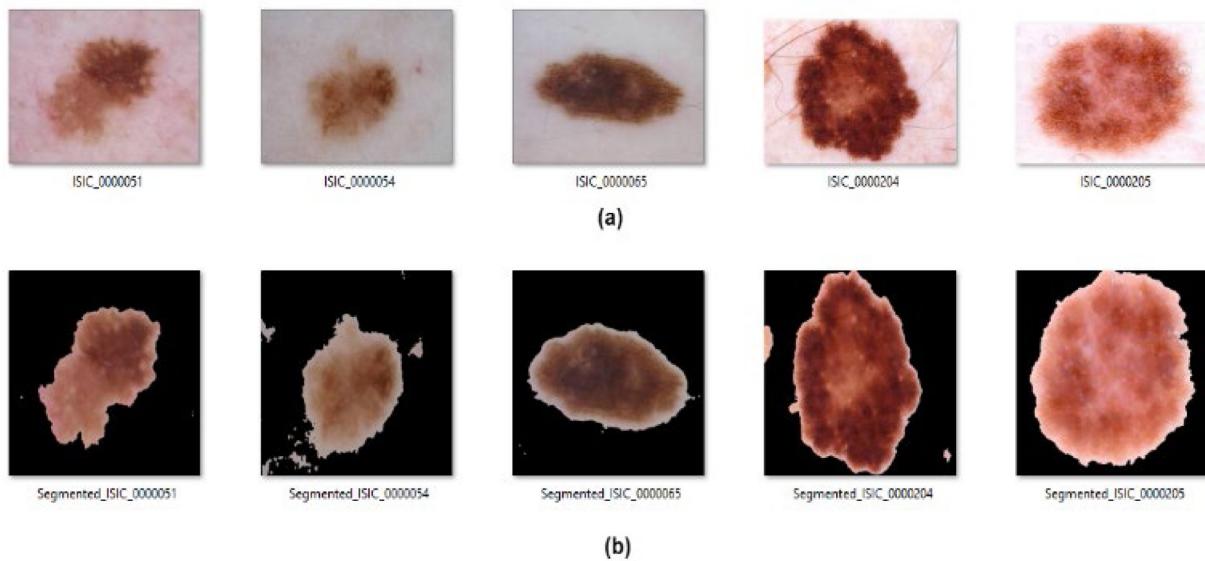


Fig. 7. Sample images before and after segmentation (a) Before segmentation (b) After segmentation.

Table 1
Different types of GLCM features.

Feature Name	Description	Formula
Energy	It yields the sum of squared elements in the GLCM with a value of 0 to 1.	$Energy = \sum_{i,j=0}^{N-1} (P_{ij})^2$
Correlation	It returns a measure of how closely a pixel is connected to its neighbors throughout the entire image.	$Correlation = \sum_{i,j=0}^{N-1} P_{ij} \frac{(i-\mu)(j-\mu)}{\sigma^2}$
Contrast	It is the overall intensity of a pixel's relationship with its neighbors.	$Contrast = \sum_{i,j=0}^{N-1} P_{ij} (i-j)^2$
Homogeneity	Homogeneity is defined as the closeness rate of distributed elements in GLCM.	$Homogeneity = \sum_{i,j=0}^{N-1} \frac{P_{ij}}{1+(i-j)^2}$
Entropy	Entropy is the degree of uniformity between pixels within the image and randomness.	$Entropy = \sum_{i,j=0}^{N-1} -\ln (P_{ij}) P_{ij}$

Where P_{ij} the GLCM, which is symmetrically normalized, N is the image's total number of gray levels, and μ is the GLCM mean and σ^2 represents the variance of the GLCM.

distinguish images as accurately as feasible. The best method to handle the region in isolation is to use the segmented lesion images to extract many features for this task. In this work, we used the GLCM features as texture features and several statistical features as color features to determine the skin disease type.

3.3.1. GLCM Features

First, every image is computed by the Gray Level Co-occurrence Matrix (GLCM). Then, the contrast, energy, entropy, correlation, and homogeneity features are calculated from the matrix. The extracted GLCM features and their description and formula are given in [Table 1](#).

3.3.2. Statistical Features

The RGB (red, green, and blue) color spaces are explored for every image, and different statistical features are extracted. This work extracted the mean, variance, standard deviation, and root mean square as the statistical features. The extracted features, along with their description and formula, are given in [Table 2](#). The extracted feature's values for some sample images are given in [Table 3](#).

3.4. Classification

The final task of the work is classification. It is the process of classifying a set of data into several categories. This work predicted the type of skin disease using the features extracted from images. Depending on the application and the given dataset's type, several

methods are employed for classification. In this work, we utilized three distinct classifiers: Support Vector Machine (SVM), K Nearest Neighbor (KNN), and Decision Tree (DT) for the categorization of skin disease. A Support Vector Machine (SVM) is a supervised classification approach used to solve classification and regression problems. It provides better accuracy in results than most of the other algorithms. To classify eight skin diseases, we used multiclass SVM that includes two approaches, the One-to-One approach and the one-to-Rest approach [44]. It applies the kernel technique to transform data that finds an optimal decision boundary between the possible outputs [45]. K-Nearest Neighbor is a non-parametric method that can be used to solve both classification and regression problems. It uses 'feature similarity' to predict unknown data points' values. In this work, the similarities between the new and available cases are found by calculating the distance between the two data points using the Euclidean distance [46]. We used the number of neighbors k=10 as it gives better classification performance. The decision tree represents the knowledge in a tree structure to make it easier to understand, which may be described as a set of discrete rules. It provides the high-speed classification of unknown records. We chose the classifier because of having the ability to use multiple feature subsets and decision rules at various stages of classification.

4. Dataset

4.1. Dataset Description

We used two well-known datasets: the "International Skin Imaging Collaboration (ISIC) 2019" challenge dataset and "HAM10000

Table 2
Different types of Statistical features.

Feature name	Description	Formula
Mean	The mean of an image is defined as the average color value in the image.	$Mean = \sum_{i,j=0}^{N-1} i P_{ij}$
Variance	The variance of an image is defined as a measure of the dispersion of the values around the mean.	$Variance = \sum_{i,j=0}^{N-1} P_{ij} (i - \mu)^2$
Standard Deviation	The standard deviation is represented as the square root of the distribution variance.	$SD = \sqrt{\sum_{i,j=0}^{N-1} P_{ij} (i - \mu)^2}$
Root Mean Square	The RMS is defined as the square root of the average of all squared intensity values.	$RMS = \sqrt{\sum_{i,j=0}^{N-1} P_{ij} (i - c)^2}$

Where, C is the Correlation factor.

Table 3
Different Features Value.

Image	Energy	Correlation	Contrast	Homogeneity	Entropy	Mean	Variance	SD	RMS
ISIC_0000000	0.44373	0.99365	0.11071	0.9939	3.7487	103.96	7214.2	101.6	11.065
ISIC_0000001	0.83246	0.96824	0.034583	0.99592	1.1987	7.8483	616.09	27.302	2.8532
ISIC_0000002	0.2976	0.96787	0.2855	0.97348	6.0337	122.31	4144.6	73.75	13.975
ISIC_0000003	0.26206	0.97325	0.30339	0.98302	5.0301	79.596	5365.9	83.485	10.98
ISIC_0000004	0.64136	0.94263	0.34138	0.98336	3.9736	31.193	2386.5	59.564	6.5033
ISIC_0000005	0.85844	0.9556	0.026747	0.99449	1.1477	6.0631	375.42	21.528	2.7896
ISIC_0000006	0.41169	0.96419	0.2384	0.97117	5.9232	132.78	2815.7	59.86	14.887
ISIC_0000007	0.46807	0.97736	0.10613	0.98848	3.4053	36.614	1915.6	55.226	6.8922
ISIC_0000008	0.57826	0.98045	0.067045	0.99159	2.6254	26.651	1641.9	48.003	6.1815
ISIC_0000009	0.82017	0.96808	0.047762	0.99484	1.1386	9.8793	741.75	31.467	2.8396

(Human-Against-Machine with 10000 training images)" dataset to train, test, and evaluate the proposed model.

4.1.1. ISIC 2019 Challenge Dataset

The ISIC is an international collection of dermatoscopic images [47]. It has developed for clinical practice as well as for supporting technical challenges. We have only used the training dataset of the ISIC 2019 challenge which contains 25331 dermoscopic images of eight different classes in the dataset: "actinic keratosis (AK), basal cell carcinoma (BCC), benign keratosis (BKL), dermatofibroma (DF), melanoma (MEL), melanocytic nevus (NV), squamous cell carcinoma (SCC) and vascular lesion (VASC)". The dataset comprises the HAM10000 (Human-Against-Machine with 10000 training images) [48] and the BCN20000 [49] dataset.

4.1.2. HAM10000 Dataset

The HAM10000 dataset consists of 10015 dermatoscopic images of pigmented skin lesions. It was collected from the Australian and Austrian patients. The images are of size 600×450 and center-cropped. The dataset comprises seven different classes: "actinic keratosis (AK), basal cell carcinoma (BCC), benign keratosis (BKL), dermatofibroma (DF), melanoma (MEL), melanocytic nevus (NV) and vascular lesion (VASC)".

4.1.3. Dataset Preparation

The ISIC 2019 challenge training dataset consists of 25331 images with eight different categories: "basal cell carcinoma (BCC), actinic keratosis (AK), melanoma (MEL), melanocytic nevus (NV), benign keratosis (BKL), dermatofibroma (DF), vascular lesion (VASC), and squamous cell (SCC)". It is an imbalanced dataset that consists of 867 images of class AK, 3323 images of class BCC, 2624 images of class BKL, 239 images of class DF, 4522 images of class MEL, 12875 images of class NV, 628 images of class SCC, and 253 images of class VASC. The HAM10000 dataset is also an imbalanced dataset where AK class has 327 images, BCC class has 514 images, BKL class has 1099 images, DF class has 115 images, MEL class has 1113 images, NV class has 6705 images and VASC class has 142 images. The data distribution of each class for ISIC 2019 dataset is given in Fig. 8. While training the model with the imbalanced dataset, there exists a bias predicting the majority classes and often ignores the minority classes. As a result, the error



Fig. 8. The distribution of the total number of images into eight classes.

increases for the minority classes and decreases for the majority classes. We applied data balancing Random Oversampling method to resolve the imbalanced dataset for both datasets.

4.1.4. Random Oversampling

Oversampling is the technique that randomly duplicates samples from minority classes. The smaller class than any class is considered a minority class. The process is then repeated until no class has a sample size smaller than the largest. Finally, the sampling is done when the balance is achieved. We have applied the oversampling method for both the ISIC 2019 and the HAM10000 datasets. Here, we only showed the distribution of the balanced ISIC 2019 dataset using oversampling method which is given in Fig. 9.

5. Results and Discussion

5.1. Experiment

We have implemented all the tasks in the python platform. We have applied the model to the ISIC 2019 challenge training dataset only containing 25331 images of eight different classes and also used the HAM10000 dataset to prove the supremacy of the proposed methods. We have completed the preprocessing task before the feature extraction.

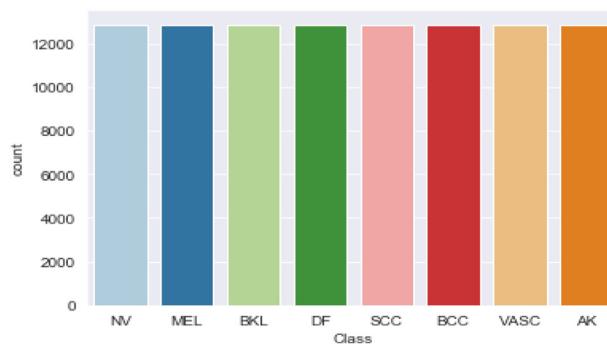


Fig. 9. Distribution of data after random Oversampling.

Table 4
Statistical measures after applying Digital Hair Removal technique.

Image	RMSE	PSNR	SSIM	SAM	UIQ
ISIC_0000043	24.85	20.22	0.72	0.13	0.98
ISIC_0000103	10.26	27.91	0.85	0.06	0.99
ISIC_0000115	17.80	23.12	0.63	0.11	0.98
ISIC_0000218	12.40	26.26	0.82	0.07	0.99
ISIC_0000219	20.39	21.94	0.76	0.11	0.99

Table 5
Statistical measures after applying Gaussian filtering.

Image	RMSE	PSNR	SSIM	SAM	UIQ
ISIC_0000096	2.93	38.77	0.92	0.02	0.99
ISIC_0000103	3.28	37.79	0.91	0.02	0.99
ISIC_0000121	2.19	41.31	0.95	0.01	0.99
ISIC_0000124	3.44	37.39	0.90	0.02	0.99
ISIC_0000125	3.06	38.41	0.90	0.02	0.99

In the first preprocessing step, we resize the input image into a size of 512×512 . It shortens the processing time as well as improves the overall performance of the system. In the second step, we have used a digital hair removal (DHR) algorithm based on morphological filterings such as Black-Hat transformation and image inpainting algorithm to remove hairs. Finally, we have used Gaussian filtering to remove noise from the input images with a 7×7 kernel where the sigma value is calculated. We have calculated some statistical measures such as Root Mean Square Error (RMSE), Peak Signal-to-Noise Ratio (PSNR), Structural Similar Index Measure (SSIM), Spectral Angle Mapper (SAM), and Universal Image Quality Index (UIQ) for the images after applying the Digital Hair Removal technique as well as the Gaussian filtering. The statistical measures are shown in Tables 4 and 5 for the images of Figs. 4 and 5 after applying the digital hair removal technique and the Gaussian filtering, respectively.

After the preprocessing task, we have applied a well-known segmentation technique as k-means and automatic Grabcut segmentation algorithms to extract the foreground of the skin diseases. However, the Grabcut method outperforms the k-means segmentation algorithm for detecting the lesion of the skin images. After the segmentation, still there are some noises in the images, and we have applied the Gaussian filtering technique again to remove the noises.

After the segmentation task, we have extracted the GLCM and the statistical features from the segmented images and formed a feature vector. After that, we applied the SVM, the KNN, and the DT algorithm to classify skin diseases. Next, we split the skin image data into two parts, 80% and 20% for training and testing, respectively. Finally, the training data parts are used to build the model, and the testing set is used to evaluate the model's performance. We also illustrated a comparative analysis of the performance of the proposed classification algorithms on the ISIC 2019 challenge dataset as well as on the HAM10000 dataset.

5.2. Performance Metrics

The classification performance of the proposed working models is evaluated by measuring some metrics like accuracy, precision, recall, F1 score, and categorical cross-entropy or log loss. The following formulas are applied to calculate these evaluation metrics.

Accuracy: Accuracy is defined as the proportion of properly-recognized samples and the total number of samples.

$$\text{Accuracy} = \left[\frac{(TP + TN)}{(TP + FN + FP + TN)} \right] \times 100\% \quad (4)$$

Precision: Precision is defined as the proportion of the properly-recognized positive samples and the total number of predicted positive samples.

$$\text{Precision} = \left[\frac{TP}{TP + FP} \right] \times 100\% \quad (5)$$

Recall: The recall is defined as the proportion of the total number of properly-recognized positive samples classified correctly.

$$\text{Recall} = \left[\frac{TP}{TP + FN} \right] \times 100\% \quad (6)$$

F1 Score: The F1 score is the weighted average that returns a single value by combining precision and recall.

$$\text{F1-Score} = 2 \times \left[\frac{(\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \right] \times 100\% \quad (7)$$

Where TP represents the true positive, the true negative is represented by TN, FP represents the false positive, and FN represents the false negative.

We have also illustrated the classification performance of our model graphically using the Receiver Operating Characteristic (ROC) and the Area Under Curve (AUC).

ROC Curve: ROC curve is the graphical illustration of the classification performance of a model. The true positive rate (TPR) is compared against the false positive rate (FPR) to plot the ROC curve.

The TPR and FPR are computed using the following formulas:

$$\text{TPR} = \frac{TP}{TP + FN} \quad (8)$$

$$\text{FPR} = \frac{FP}{FP + TN} \quad (9)$$

AUC: AUC illustrates the whole two-dimensional region beneath the ROC curve, from point (0, 0) to point (1, 1). The greater the AUC, the better the model's performance.

Categorical Cross-Entropy: Categorical cross-entropy or log loss is a metric to measure the performance of the classification model whose value ranges from 0 to 1. It is a loss function, used in multi-class classification tasks. In this task, the model must determine an example can only belong to one out of many possible categories. Formally, it is designed to quantify the difference between two probability distributions. In general, minimizing Categorical cross-entropy increases the classifier's accuracy. If there are N samples belonging to M classes, then the Categorical Cross-Entropy is calculated as follows:

$$\log loss = - \frac{1}{N} \sum_i^N \sum_j^M y_{ij} \log(p_{ij}) \quad (10)$$

Where N is the number of rows or samples, M is the number of classes, y_{ij} is 1 if the sample i belongs to class j else 0, and p_{ij} is the probability of our classifier that predicts sample i to class j .

5.3. Result and Analysis

In this section, we have evaluated the performance of our proposed classification model and compared it with the performance of existing techniques. As our datasets are imbalanced, we have evaluated our model using the imbalanced datasets as well as after balancing the datasets.

****Classification Report - Support Vector Machine (SVM) ****					
	precision	recall	f1-score	support	
AK	0.00	0.00	0.00	66	
BCC	0.23	0.06	0.10	97	
BKL	0.28	0.08	0.12	225	
DF	0.00	0.00	0.00	21	
MEL	0.61	0.60	0.61	212	
NV	0.74	0.95	0.83	1313	
VASC	0.50	0.03	0.06	32	
accuracy			0.71	1966	
macro avg	0.34	0.24	0.24	1966	
weighted avg	0.61	0.71	0.64	1966	

Fig. 10. Classification Report for the HAM10000 dataset using SVM classifier while the dataset is imbalanced.

Table 6

Performance metrics of our proposed models for Imbalanced datasets (Before applying the Oversampling technique on the datasets).

Dataset	Model	Accuracy	Precision	Recall	F1-score
ISIC 2019	SVM	52.00	20.63	18.63	18.38
	KNN	42.00	18.63	18.75	18.63
	DT	40.00	17.88	17.75	17.75
HAM10000	SVM	71.00	33.71	24.57	24.57
	KNN	55.00	23.00	21.71	22.71
	DT	57.00	22.14	24.29	23.86

5.3.1. Result for the Imbalanced Datasets

The overall classification performance before data balancing for the Support Vector Machine, K-Nearest Neighbor, and Decision Tree classifiers using the ISIC 2019 dataset and the HAM10000 dataset are illustrated in Table 6. For details, we have also illustrated the disease-wise classification report for the SVM classifier using the HAM10000 dataset as shown in Fig. 10. We found that the majority class NV was predicted as well where the minority class DF was not. This is due to the imbalanced dataset problem.

5.3.2. Result after Balancing the Datasets

Firstly, we illustrated the classification performance of the proposed work for the three different classification algorithms and compared them with each other. Then, it explains the model's improvement on sensitivity results using the SVM algorithm (95%) in comparison to the decision tree (93%) and even for the case of the KNN (94%) algorithm, which indicates that the model outperforms on the large dataset using the SVM algorithm, as shown in Fig. 11. and Table 7. Nevertheless, the average precision and average F1 score values obtained using the SVM algorithm (95.13% and 94.88%) are also greater. Table 7 also shows that the SVM classifier performs better with an average accuracy of 95% and 97% among the three classifiers for both datasets, respectively. We have also evaluated our proposed model using the HAM10000 dataset and found that our model also performs well for that dataset, as shown in Fig. 12 and Table 7. From the Table 7, we have also observed that Support Vector Machine classifier has a small number of log loss for both datasets which indicates the better classification accuracy.

The average performance obtained by the SVM classifier is more excellent than the other two classifiers which is shown in Fig. 11.

The average performance obtained by the SVM classifier is more excellent than the other two classifiers which is shown in Fig. 12.

In addition, the confusion matrices given in Figs. 13 and 14 show the classification performance of the investigated algorithms. In the confusion matrix, the value in each row represents the corresponding actual labels, and the value in each row represents the corresponding predicted labels. Thus, the cell value depicts the percentage of prediction. At the same time, the diagonal cell shows the highest level of

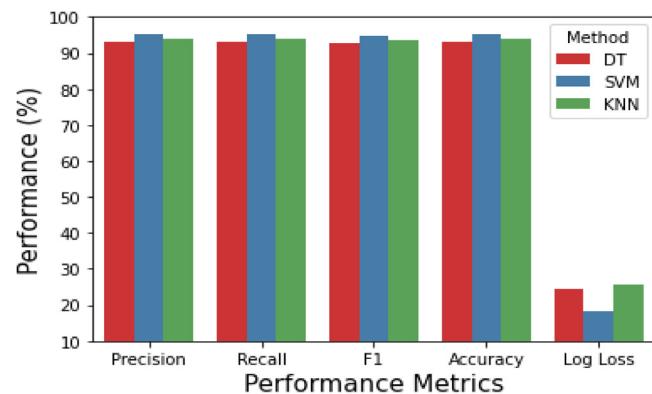


Fig. 11. Comparison among three classifiers based on the Precision, Recall, F1, Accuracy, and Log Loss values for ISIC 2019 dataset.

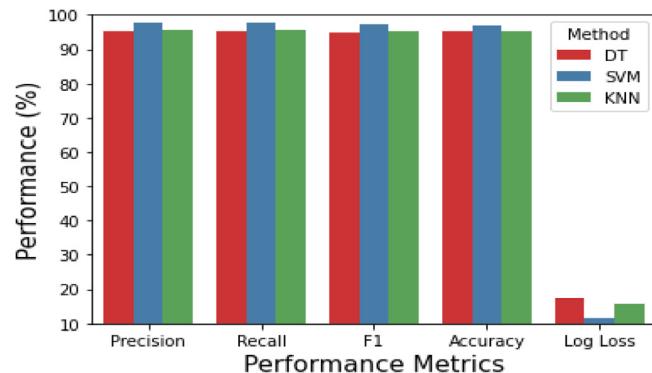


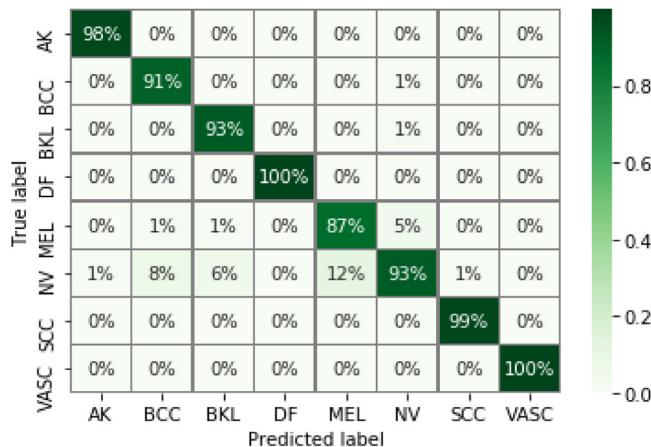
Fig. 12. Comparison among three classifiers based on the Precision, Recall, F1, Accuracy, and Log Loss values for HAM10000 dataset.

predictions in the confusion matrix that indicates a minimum error rate for each class of skin diseases. Using the confusion matrix, we computed the accuracy, precision, recall, and f1-score to test the proposed model's performance.

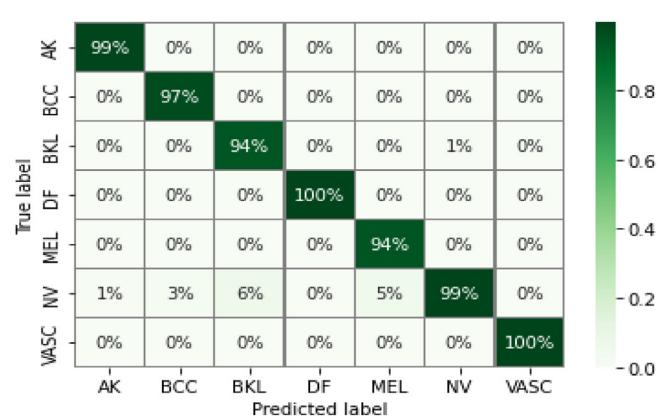
To solve our research problems, we also evaluated the classification performance of our proposed models for each type of skin disease by illustrating the results of the ROC analysis shown in Fig. 15 (a, b, c), and Fig. 16 (a, b, c). Again, each classifier showed similar behavior. Here, we represented the type of disease for ISIC 2019 dataset as follows: class 0 for Actinic Keratosis (AK), class 1 for Basal Cell Carcinoma (BCC), class 2 for Benign Keratosis (BKL), class 3 for Dermatofibroma (DF), class 4 for Melanoma (MEL), class 5 for Melanocytic Nevus (NV), class 6 for Squamous Cell Carcinoma (SCC), and class 7 for Vascular

Table 7
Performance metrics of our proposed models for balanced dataset.

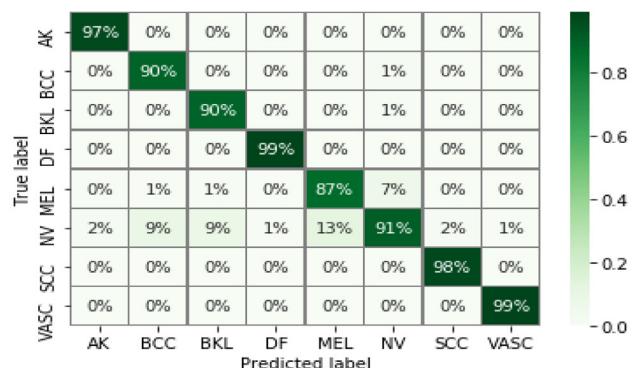
Dataset	Model	Accuracy	Precision	Recall	F1-score	Log loss (%)
ISIC 2019	SVM	95.00	95.13	95.00	94.88	18.09
	KNN	94.00	93.88	93.88	93.38	25.49
	DT	93.00	93.13	93.00	92.50	24.49
HAM10000	SVM	97.00	97.71	97.57	97.43	11.37
	KNN	95.00	95.71	95.57	95.14	15.59
	DT	95.00	95.14	95.14	94.71	17.37



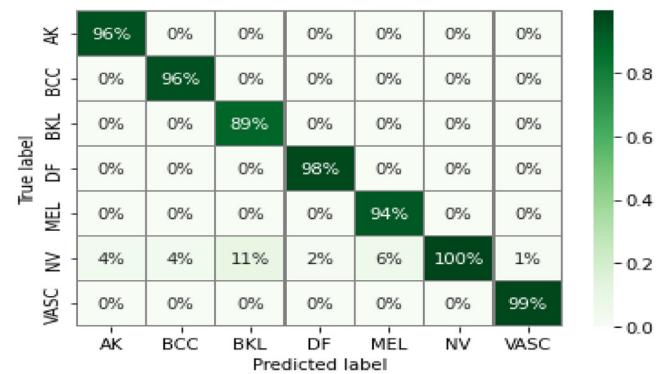
(a)



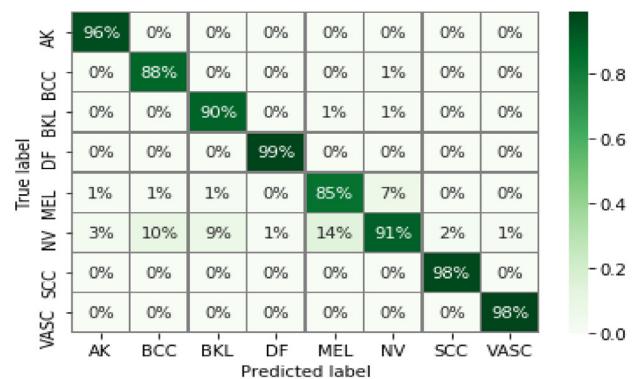
(a)



(b)



(b)



(c)

Fig. 13. The confusion matrix for the ISIC 2019 dataset of the proposed system using the (a) SVM classifier, (b) KNN classifier, and (c) Decision Tree classifier, which represent the predicted values of the eight classes in percentage.

Fig. 14. The confusion matrix for the HAM10000 dataset of the proposed system using the (a) SVM classifier, (b) KNN classifier, and (c) Decision Tree classifier, which represent the predicted values of the eight classes in percentage.

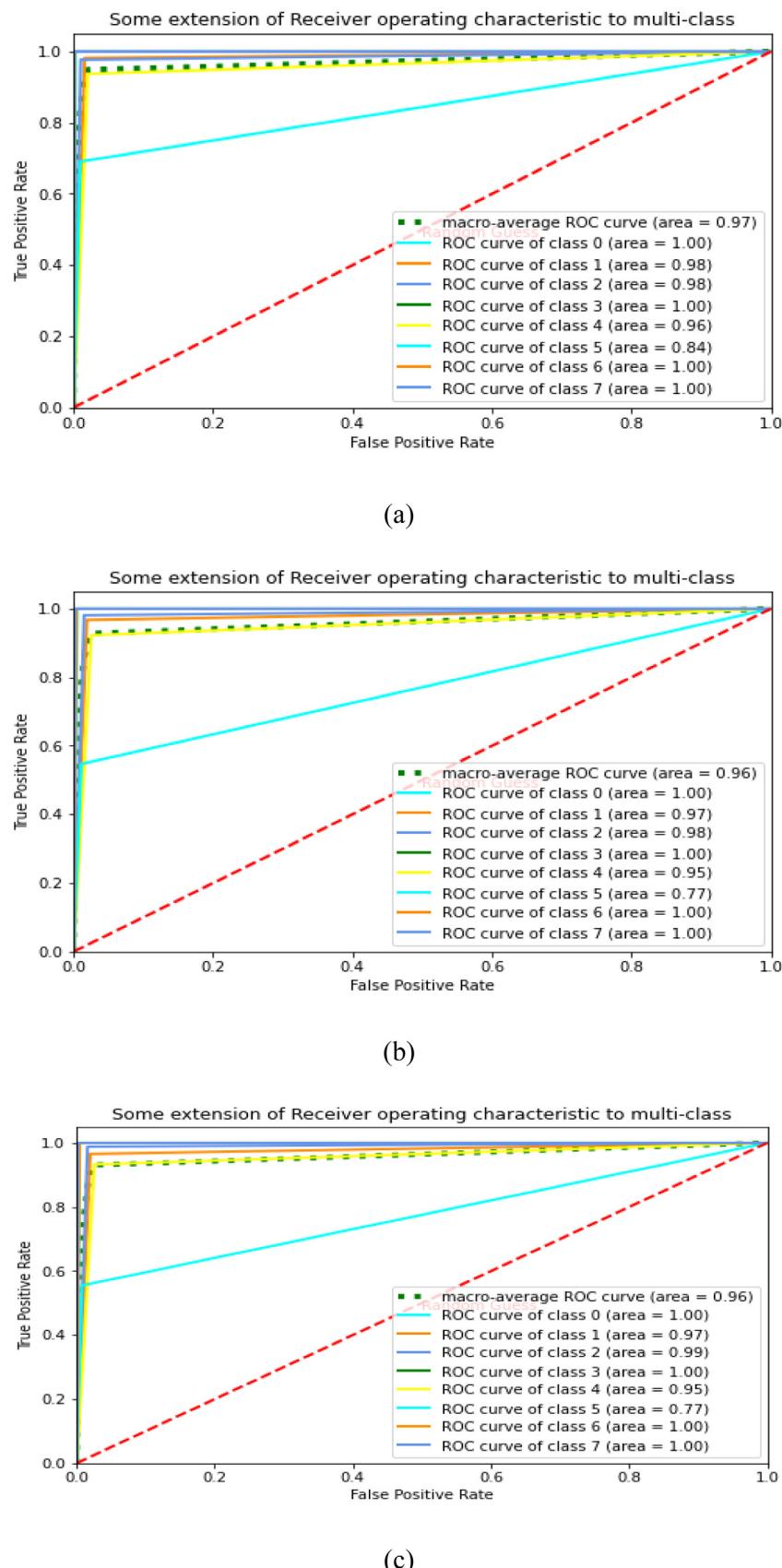


Fig. 15. ROC performance curves for the ISIC 2019 dataset of the proposed classification model using the (a) SVM classifier, (b) KNN classifier, and (c) Decision Tree Classifier. The curves close to the top left corner indicate the classification accuracy of each class in the SVM classifier. The area Under the Curve (AUC) shows the model's ability to separate diseases.

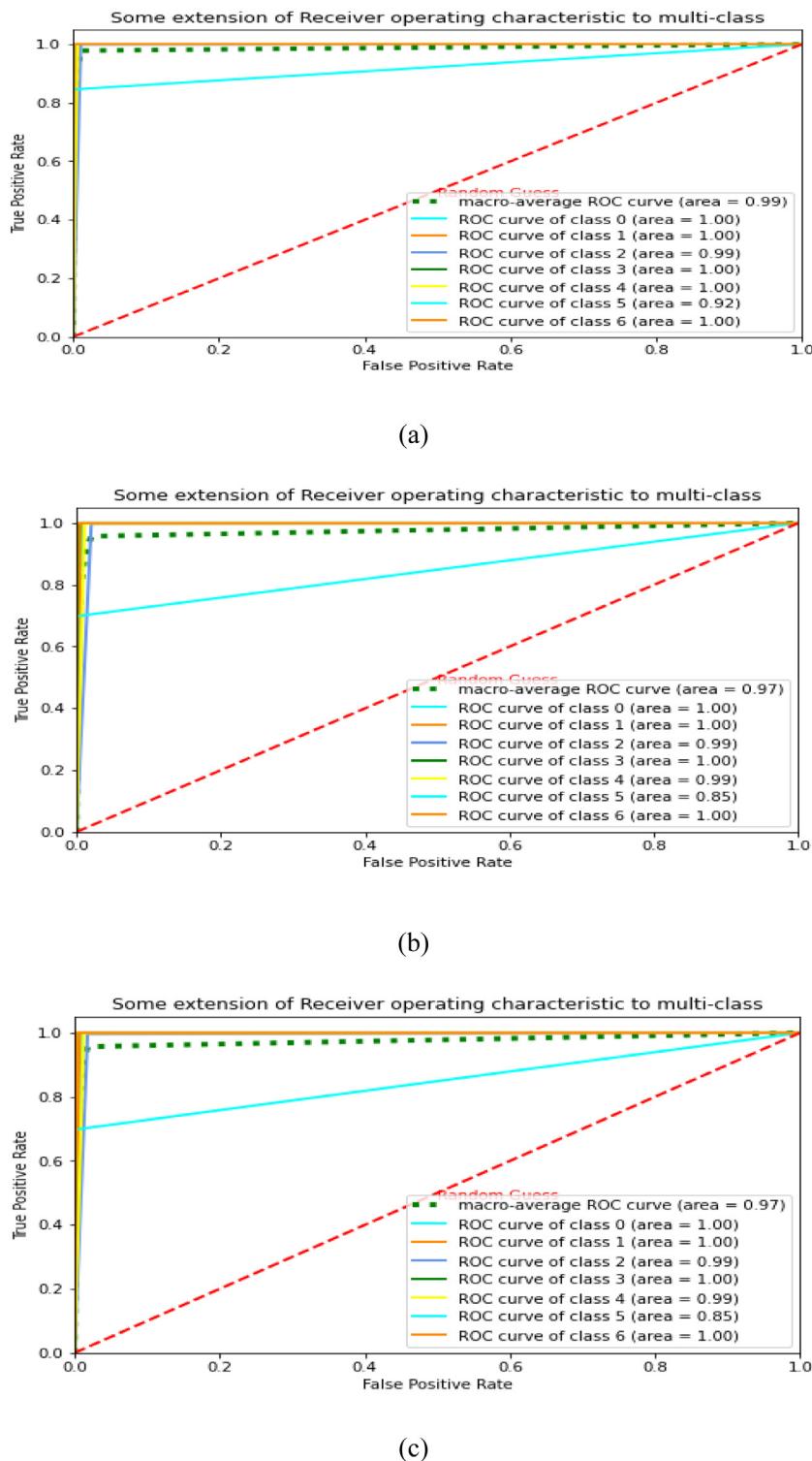


Fig. 16. ROC performance curves for the HAM10000 dataset of the proposed classification model using the (a) SVM classifier, (b) KNN classifier, and (c) Decision Tree Classifier. The curves close to the top left corner indicate the classification accuracy of each class in the SVM classifier. The area Under the Curve (AUC) shows the model's ability to separate diseases.

Lesion (VASC). We achieved the highest classification performance for class 0, class 3, class 6, and class 7 using the three different classifiers. However, class 5 achieved the minimum AUC. The overall AUC of this system was 97% for SVM classifier. The AUC shows better detection using the SVM than the DT and the KNN. For the HAM10000 dataset, the type of disease in the ROC is represented as follows: class 0 for

Actinic Keratosis (AK), class 1 for Basal Cell Carcinoma (BCC), class 2 for Benign Keratosis (BKL), class 3 for Dermatofibroma (DF), class 4 for Melanoma (MEL), class 5 for Melanocytic Nevus (NV), class 6 for Vascular Lesion (VASC) as shown in Fig. 15. We have also achieved the highest classification performance for class 0, class 1, class 3, class 4, and class 6 using the three different classifiers. However, class 5

Table 8

Comparative analysis of our proposed model with State-of-the-art methods.

Model	Dataset	Classifier	Feature	Segmentation	Accuracy
Proposed method	ISIC-2019	SVM, KNN, DT	GLCM + Statistical	Automatic Grabcut	95%, 94%, 93%
	HAM10000				97%, 95% 95%
Kethana et al. 2022 [16]	ISIC-2019	CNN	-	-	92%
Maduranga et al. 2022 [19]	HAM10000	CNN (MobileNet)	-	-	85%
Jain et al. 2022 [20]	ISIC	OP-DNN	Color Feature + Texture Feature	-	95%
Janney et al. 2018 [21]	ISIC	SVM	GLCM + Color + ABCD	Manual	71%
Albawi et al. 2019 [32]	ISIC	SVM, KNN	2D-DWT + GLCM	Region Growing	91.13%, 87.46%
Sinthura et al. 2020 [24]	Other datasets	SVM	GLCM	Otsu's Method	89%
Ubale et al. 2019 [31]	Other datasets	KNN	HSV + LAB	-	91.80%
Ozkan et al. 2017 [33]	Other datasets	SVM, KNN, DT	ABCD	-	89.50%, 82%, 90%

achieved the minimum AUC. The overall AUC of this system was 99% for SVM classifier. The AUC shows better detection using the SVM than other methods: DT and KNN.

5.4. Comparison with State-of-the-art Methods

We have compared our result with some well-established recent methods including deep learning [12], [19–22, 24, 31, 32], and [33] and shown in Table 8. Our technique outperformed these state-of-the-art methods.

Kethana et al. [16], Maduranga et al. [19] and Jain et al. [20] all have used deep learning methods (CNN, OP-DNN). Complexity of a convolutional neural network model is $O(k * n * d^2)$, where, n is the sequence length, d is the representation dimension and k is the kernel size of convolution. On the other hand, we have used SVM, KNN, DT. Each of these methods have less computational complexity than deep learning-based models. Moreover, our model has shown better accuracy than the mentioned methods.

Janney et al. [21] have used a manual method for segmenting the diseased region. On the other hand we have used an automatic segmentation method. Manual segmentation method requires more time than automatic segmentation method. Besides, we have far better accuracy than theirs.

Albawi et al. [32] have used a region growing method for segmenting the diseased region where we used automatic grabcut segmentation. The complexity of the region growing method, $O(n^2)$, where n is the number of pixels in the image, is greater than the complexity of automatic grabcut segmentation method, $O(n \log n)$, where n is the number of pixels in the image. Moreover, our model has shown better accuracy than the mentioned method.

Sinthura et al. [24] have used Otsu's method for segmenting diseased regions which is more complex, having complexity of $O(LN)$, where N is the total number of pixels in the image and L is the pixel intensity, than the automatic grabcut segmentation method. Moreover, the performance of our model is better than the mentioned method.

Ubale et al. [31] and Ozkan et al. [33] have not used any segmentation method. From this point of view they have less computational complexity than ours. But extracting features without segmenting the diseased region can have a severe negative effect on the model's performance. And this is prominent from their accuracy measure. Although our complexity is slightly more than theirs because of our having a segmentation method, we think it is worthy to have it for better performance.

6. Conclusion and Future Work

Skin disease is currently a global problem. People of many countries or regions suffer from different types of skin diseases. We can fight against these diseases by developing various techniques and processes. In this research, we have performed the work with several phases. We

applied a customized digital hair removal technique using Morphological Black-Hat Transformation to remove hairs and Gaussian Filter to blur the images. After that, we used an automatic Grabcut segmentation to detect the skin lesion, which perfectly segmented and detected the disease region. Finally, we extracted the GLCM and some statistical features and applied them to the SVM, KNN, and DT classifiers to classify the type of skin disease. We used two benchmark publicly available datasets: ISIC 2019 challenge and HAM10000. As, these datasets are somehow imbalanced, so we performed data balancing using random over sampling technique. We obtained an average accuracy of 95%, 94%, and 93% for the ISIC 2019 dataset using SVM, KNN, and DT classifiers, respectively. Similarly, we have obtained accuracy of 97%, 95%, and 95% for the HAM10000 dataset using SVM, KNN, and DT classifiers, respectively. It indicates that our model performs better using the HAM10000 dataset than the ISIC2019 dataset. We also observed that our model dramatically performs well for the balanced data. Our model outperforms some of the state-of-the-art methods for skin disease classification.

We can use this model for other skin disease classification tasks. But still, there is some scope to improve the classification performance. We used an automatic segmentation algorithm that sometimes does not detect the skin lesion correctly. As a result, it leads to misclassification, which is the limitation of our study. With the support of more efficient segmentation and classification techniques such as ensemble learning and deep learning techniques, future research will focus on real-time skin disease detection. In addition, we believe that it will increase the performance and accuracy of image classification and object detection systems algorithms. We hope it will be useful for the patients for an early detection of diseases to keep their skin healthy.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

- [1] Anatomy of the skin, Stanford children's health, 2021, [Online]. Available: <https://www.stanfordchildrens.org/en/topic/default?id=anatomy-of-the-skin-85-P01336> (Accessed 11 June 2021).
- [2] M.W. Greaves, Skin disease, Britannica, 29, 2020, [Online]. Available: <https://www.britannica.com/science/human-skin-disease>. (Accessed 11 June 2021).
- [3] Bangladesh: Skin disease, 2018, [Online]. Available: <https://www.worldlifeexpectancy.com/bangladesh-skin-disease>. (Accessed 9 July 2021).
- [4] Cancer, world health organization, 21, 2021, [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/cancer>. (Accessed 20 March 2020).
- [5] Md. Al Mamun, Mohammad Sharif Uddin, A survey on a skin disease detection system, Int. J. Healthc. Inform. Syst. Inform. 16 (4) (2021) 1–17.

- [6] C.N. Vasconcelos, B.N. Vasconcelos, Experiments using deep learning for dermoscopy image analysis, *Pattern Recognit. Lett.* 139 (2020) 95–103.
- [7] U.-O. Dorj, K.K. Lee, J.Y. Choi and. M. Lee, The skin cancer classification using deep convolutional neural network, *Multimedia Tools Appl.* 77 (2018) 9909–9924.
- [8] M. Taufiq, N. Hameed, A. Anjum, F. Hameed, m-Skin Doctor: A Mobile Enabled System for Early Melanoma Skin Cancer Detection Using Support Vector Machine, in: eHealth 360°. International Summit on eHealth, 2017, pp. 468–475.
- [9] Jagdis et.al., J.A.D.L. Cruz-Vargas, M.E.R. Camacho, Advance study of skin diseases detection using image processing methods, *Nat. Volatiles Essent. Oils J.* 9 (1) (2022) 997–1007.
- [10] Z. Naeem, G. Zia, Z. Bukhari, A healthcare model to predict skin cancer using deep extreme machine, *J. NCBAE* 1 (2) (2022) 23–30.
- [11] S.K. Bandyopadhyay, P. Bose, A. Bhaumik, S. Poddar, Machine learning and deep learning integration for skin diseases prediction, *Int. J. Eng. Trends Technol.* 70 (2) (2022) 11–18.
- [12] A. Kalaivani, S. Karpagavalli, Detection and classification of skin diseases with ensembles of deep learning networks in medical imaging, *Int. J. Health Sci.* 6 (S1) (2022) 13624–13637.
- [13] S.A. AlDera, M.T.B. Othman, A model for classification and diagnosis of skin disease using machine learning and image processing techniques, *Int. J. Adv. Comput. Sci. Appl.* 13 (5) (2022).
- [14] P.R. Kshirsagar, H. Manoharan, S. Shitharth, A.M. Alshareef, N. Albishry, P.K. Balachandran, Deep learning approaches for prognosis of automated skin disease, *Life* 2022 12 (426) (2022).
- [15] M.Q. Hatem, Skin lesion classification system using a Knearest neighbor algorithm, in: Visual Computing for Industry, Biomedicine, and Art. Vol. 5, (7) 2022.
- [16] K.S. A, M.S. B, Melanoma disease detection and classification using deep learning, *Int. J. Res. Appl. Sci. Eng. Technol.* 10 (7) (2022).
- [17] P. Yao, Single model deep learning on imbalanced small datasets for skin lesion classification, *IEEE Trans. Med. Imaging* 41 (5) (2022) 1242–1254.
- [18] K. Padmavathi, H. Neelam, M.P.K. Reddy, P. Yadlapalli, K.S. Veerella, K. Pampari, Melanoma detection using deep learning, in: 2022 International Conference on Computer Communication and Informatics, ICCCI, 2022.
- [19] M. Maduranga, D. Nandasena, Mobile-based skin disease diagnosis system using convolutional neural networks (CNN), *I.J. Image Graphics Signal Process.* 3 (2022) 47–57.
- [20] A. Jain, A.C.S. Rao, P.K. Jain, A. Abraham, Multi-type skin diseases classification using OP-DNN based feature extraction approach, *Multimedia Tools Appl.* 81 (2022) 6451–6476.
- [21] B. Janney, S.E. Roslin, M.J. Shelcy, A comparative analysis of skin cancer detection based on SVM, ANN and naive bayes classifier, in: International Conference on Recent Innovations in Electrical, Electronics & Communication Engineering, ICRIEECE, 2018.
- [22] H. Alquran, I.A. Qasmieh, A.M. Alqudah, S. Alhammouri, E. Alawneh, A. Abughazaleh, F. Hasayen, The melanoma skin cancer detection and classification using support vector machine, in: 2017 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies, AEECT, 2017.
- [23] V.R. Balaji, S.T. Suganthi, R. Rajadevi, V. Krishna Kumar, B. Saravana Balaji, Sanjeevi Pandiyar, Skin disease detection and segmentation using dynamic graph cut algorithm and classification through naive Bayes classifier, *Measurement* 163 (2020).
- [24] S.S. Sinthura, K.R. Sharon, G. Bhavani, L. Mounika, B. Joshika, Advanced skin diseases diagnosis leveraging image processing, in: 2020 International Conference on Electronics and Sustainable Communication Systems, ICESC, Coimbatore, India, 2020.
- [25] N.V. Kumar, P.V. Kumar, K. Pramodh, Y. Karuna, Classification of skin diseases using image processing and SVM, in: International Conference on Vision Towards Emerging Trends in Communication and Networking, ViTECoN, Vellore, India, 2019.
- [26] N. Hameed, A.M. Shabut, M.A. Hossain, Multi-class skin diseases classification using deep convolutional neural network and support vector machine, in: 2018 12th International Conference on Software, Knowledge, Information Management & Applications, SKIMA, Phnom Penh, Cambodia, 2018.
- [27] T. Shanthi, R.S. Sabeenian, R. Anand, Automatic diagnosis of skin diseases using convolution neural network, *Microprocess. Microsyst.* 76 (2020) 103074, 0141-9331.
- [28] R. Bhavani, V. Prakash, R. Kumaresan, R. Srinivasan, Vision-based skin disease identification using deep learning, *Int. J. Eng. Adv. Technol. (IJEAT)* 8 (6) (2019) 3784–3788.
- [29] N. Hameed, A. Shabut, M.A. Hossain, A computer-aided diagnosis system for classifying prominent skin lesions using machine learning, in: 2018 10 th Computer Science and Electronic Engineering Conference, CEEC, Colchester, UK, 2019.
- [30] N. Hameed, F. Hameed, A. Shabut, S. Khan, S. Cirstea, A. Hossain, An intelligent computer-aided scheme for classifying multiple skin lesions, *Computers* 62 (2019).
- [31] A. Ubale, P. Paikrao, Detection and classification of skin diseases using different ColorPhase models, *Int. Res. J. Eng. Technol.* 6 (7) (2019) 1331–1335.
- [32] S. Albawi, Y. Almadanie, Y.A. Abbas, Robust skin diseases detection and classification using deep neural networks, *Int. J. Eng. Technol.* 7 (4) (2019) 6473–6480.
- [33] I.A. Ozkan, M. Koklu, Skin lesion classification using machine learning algorithms, *Int. J. Intell. Syst. Appl. Eng.* 5 (2017) 285–289.
- [34] G. Capdehourat, A. Corez, A. Bazzano, R. Alonso-Suárez, Toward a combined tool to assist dermatologists in melanoma detection from dermoscopic images of pigmented skin lesions, *Pattern Recognit. Lett.* 32 (2011) 2187–2196.
- [35] D. Ruiz, V. Berenguer, A. Soriano, B. Sánchez, A decision support system for the diagnosis of melanoma: A comparative approach, *Expert Syst. Appl. Int. J.* 38 (12) (2011) 15217–15223.
- [36] A. Nasonova, A. Nasonov, A. Krylov, I. Pechenko, A. Umnov, N. Makhneva, Image warping in dermatological image hair removal, in: 11th International Conference on Image Analysis and Recognition, ICIAR '14, Vilamoura, Portugal, 2014.
- [37] R. Garnavi, M. Aldeen, M.E. Celebi, A. Bhuiyan, C. Dolianitis, G. Varigos, Automatic segmentation of dermoscopy images using histogram thresholding on optimal color channels, *Int. J. Med. Med. Sci.* 1 (2010) 126–134.
- [38] S. Sookpotharam, Border detection of skin lesion images based on fuzzy C-means thresholding, in: WGEC '09: Proceedings of the 2009 Third International Conference on Genetic and Evolutionary Computing, 2009.
- [39] A. Swain, Noise filtering in digital image processing, medium, 3, 2018, [Online]. Available: <https://medium.com/image-vision/noise-filtering-in-digital-image-processing-d12b5266847c>. (Accessed 20 December 2020).
- [40] K. Bhargavi, S. Jyothi, A survey on threshold based segmentation technique in image processing, *Int. J. Innov. Res. Dev.* 3 (12) (2014) 234–239.
- [41] H.G. Kaganami, B. Zou, Region-based segmentation versus edge detection, in: 2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Kyoto, Japan, 2009.
- [42] A. Masood, A.A. Al-Jumaily, Computer aided diagnostic support system for skin cancer: A review of techniques and algorithms, *Int. J. Biomed. Imag.* 22 (2013).
- [43] L. Barghout, J. Sheynin, Real-world scene perception and perceptual organization: Lessons from computer vision, *J. Vis.* 13 (2013).
- [44] Multiclass classification using support vector machines, Baeldung, 25, 2021, [Online]. Available: <https://www.baeldung.com/cs/svm-multiclass-classification>. (Accessed 30 August 2021).
- [45] A. Goel, S.K. Srivastava, Role of kernel parameters in performance evaluation of SVM, in: 2016 Second International Conference on Computational Intelligence & Communication Technology, CICT, 2016.
- [46] M. Chatterjee, A quick introduction to KNN algorithm, 3, 2020, [Online]. Available: <https://www.mygreatlearning.com/blog/knn-algorithm-introduction>. (Accessed 28 June 2021).
- [47] N. Codella, D. Gutman, B. Helba, M.E. Celebi, M. Combalia, J.M.H. Kittler, V. Rotemberg, P. Tschandl, Skin lesion analysis towards melanoma detection, in: ISIC, 2019, [Online]. Available: <https://challenge2019.isic-archive.com>. (Accessed 2 February 2020).
- [48] P. Tschandl, C. Rosendahl, H. Kittler, The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions, *Sci. Data* 14 (2018).
- [49] M. Combalia, N.C.F. Codella, V. Rotemberg, B. Helba, V. Vilaplana, O. Reiter, C. Carrera, A. Barreiro, A.C. Halpern, S. Puig, J. Malvehy, BCN20000: dermoscopic lesions in the wild, 2019.