Received: 24 October 2022 | Revised: 5 December 2022 | Accepted: 12 December 2022 | Published online: 14 December 2022

RESEARCH ARTICLE

Journal of Computational and Cognitive Engineering
2023, Vol. 2(3) 226–235
DOI: 10.47852/bonviewJCCE2202482

BON VIEW PUBLISHING

Efficient and Low-Cost Skin Cancer Detection System Implementation with a Comparative Study Between Traditional and CNN-Based Models

Lakindu Induwara Mampitiya[1,*], Namal Rathnayake[2] and Subashini De Silva[1]

[1]*Department of Electrical and Electronic Engineering, Sri Lanka Institute of Information Technology, Sri Lanka*

[2]*School of Systems Engineering, Kochi University of Technology, Japan*

Abstract: Medical image classification is an essential task in the field of combining medical applications with Artificial Intelligence. This study is carried out to introduce an accurate, precise method for skin cancer recognition. This research investigates the performance of classifying skin cancer dataset HAM10000 using ResNet50, MobileNet, and the traditional support vector machine (SVM) model. The dataset combines seven cancer types: actinic keratosis, basal cell carcinoma, benign keratosis, dermatofibroma, melanoma, melanocytic nevus, and vascular lesion. The SVM classifier is designed to employ a histogram of oriented gradient (HOG) features with principle component analysis (PCA). Moreover, the Synthetic Minority Oversampling Technique is used to balance the dataset. Additionally, six conventional machine learning (ML) methods are used to compare the results with the calculation of precision, recall, F1-score, and accuracy. The results confirm that the SVM method outperforms the other algorithms with an accuracy of 99.15%. The novelty contribution of this research activity is mainly based on the development of a high accuracy, low computational complex machine method for skin cancer types recognition in the domain of medical image classification.

Keywords: HOG, MobileNet, PCA, ResNet50, skin cancer, SMOTE, SVM

1. Introduction

According to research studies on skin cancer by the World Health Organization (WHO), globally, there are between 2 and 3 million nonmelanoma skin cancer patients, and 132,000 melanoma skin cancer patients are identified yearly. Moreover, it has been determined that melanoma is in the 17th place on the world cancer list. The main reasons are the depletion of the ozone layer, more UV radiation reaching the Earth's surface, and the consumption of arsenic-contained water (World Health Organization, 2022).

Furthermore, the following individual factors can cause skin cancers: clear skin, severe sunburns, genetic history, many moles, blue, green, or hazel eyes, and radiation exposure.

Medical image classification is one of the main essential fields in medical studies. The predictions about the patient's disease severity and the medications that should be given to the patients depend on the results given out by the medical image classification models. According to dermatologists, if skin cancers can be identified at the initial stage of growing cancer, there is a chance to eliminate cancer with the proper treatments. Therefore, this research was carried out to develop an accurate skin cancer recognition system by analyzing the histogram of oriented gradients (HOG) features. The study combines feature extraction methods, dimension reduction techniques, oversampling data, and machine learning (ML) techniques to carry out a higher-performance model to identify the seven common types of skin cancers.

*Corresponding author: Lakindu Induwara Mampitiya, Department of Electrical and Electronic Engineering, Sri Lanka Institute of Information Technology, Sri Lanka. Email: lakinduinduwara21@gmail.com

2. Literature Review

There have been numerous studies have been carried out to identify early skin cancers based on various kinds of methods.

Moldovan (2019) has carried out research in the field of image-based skin cancer recognition. The proposed solution is a two-step recognition method in which step 1 produced an accuracy of 85.0%, and in step 2 of the proposed method, 75.0% accuracy was achieved. The research uses the Human Against Machine with 10000 training images (HAM10000) dataset. Shah (2021) explains the usage of the LRNet, a deep convolution neural network (DCNN), to classify skin cancer images with low resolution. The research is carried out based on the publicly available HAM10000 dataset. This study has achieved an accuracy score of 90.6%. Moreover, it has reached a precision of 94.2% and 94.0% sensitivity. Waweru et al. (2020) conducted a DCNN to identify skin cancer using the HAM10000 publicly available dataset, the training dataset for the proposed model

where the seven skin cancer types are described. The proposed method in this research achieved a balanced accuracy of 78.0%. The melanoma detection model developed by Cakmak and Tenekeci (2021) using the Basnet mobile neural network achieved an accuracy of 89.20% with the usage of the HAM10000 imbalanced dataset. Moreover, with the balancing of the dataset, 97.90% accuracy was achieved in skin cancer recognition. Huo (2021) developed a convolution neural network (CNN) model that achieved an accuracy of 75% for the classification of skin cancer images. In this study, the researcher has experimented with three models: the CNN classification model, CNN, categorical multi-layer perceptron (MLP), and CNN combined with categorical and numerical MLP.

Mainly, a set of skin cancers and skin cancers are considered. Subha et al. (2020) research discusses skin cancer detection and differentiation here. For the classification process, this research study has used CNN, where the attained maximum accuracy is 80.2%.

Demir et al. (2019) have conducted an early-detection skin cancer methodology considering two diseases. Malignant and benign diseases are considered, and classification approaches are designed according to those. ResNet-101 architecture and Inception-v3 architecture are carried out in this study. These approaches successfully attained accuracies of 84.09% and 87.42%.

According to the past studies on the mentioned skin cancer dataset, it is noticeable that there is still a possibility of increasing the classification accuracy. Therefore, the current research is carried out to investigate a method with higher accuracy and lower weight in computation.

3. Methodology

This section introduces the related studies on methodology and the experimental platform, such as dataset preparation, image preprocessing, feature extraction, data oversampling, feature dimension reduction, and development of the classification models.

When considering the research studies mentioned in the literature review, it can be identified that most of the models achieved lower accuracies due to not considering the data imbalance and the noise that can interrupt the whole model from the background skin other than focusing on the cancer-affected area of skin.

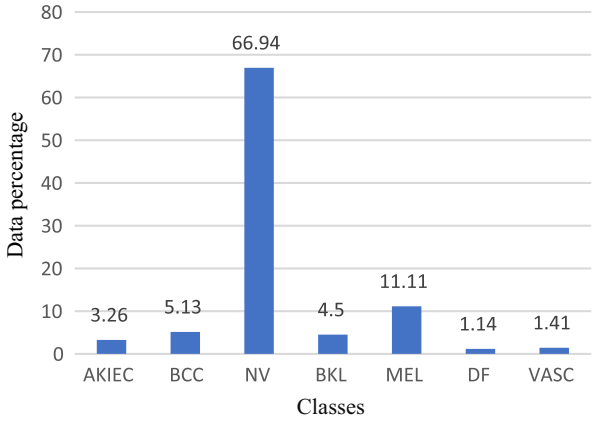
3.1. Dataset

Human Against Machine with 10000 (HAM10000) training images was used as the dataset of this study which is a publicly available dataset on skin cancers assembled with 10015 images (Tschandl et al., 2018). This dataset contains RGB images of skins with a size of 450 × 600. The sample size of the dataset is 10015 images that are collected from different populations.

The dataset describes seven common diseases that can be identified with the visible eye. More than 50% of the dataset images are confirmed by pathology. The dataset mainly represents the following skin cancers with different sample sizes: actinic keratosis (AKIEC), basal cell carcinoma (BCC), benign keratosis (BKL), dermatofibroma (DF), melanoma (MEL), melanocytic nevus (NV), and vascular lesion (VASC). Since the sample size of each disease mentioned above is different, the dataset is imbalanced.

Figure 1 illustrates the imbalanced distribution of the HAM10000 dataset. The melanocytic nevus (NV) skin cancer type includes 6705 images and is the class with the highest number of samples. In comparison, more than 50% of the dataset images belong to class NV, and the class with the least number of images can be identified as DF, which is composed of 115 images having closer to 1% of the dataset.

Figure 1
Statistical representation of the dataset



3.2. Image preprocessing

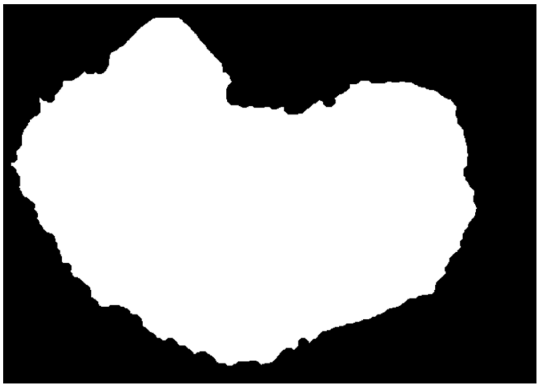
In skin cancer recognition processes, it is crucial to focus only on cancer-affected areas of skin in which the whole model depends on the quality and cleanness of the obtained images. Since there are different kinds of skin types and different kinds of skin pigments, focusing on the skin color patch is more important. Moreover, the noises in the images can negatively impact the accuracy and effectiveness of the model. Furthermore, human hair appearing or covering the affected area also negatively impacts the accuracy of the skin cancer recognition model. As the cleanness of the images is reduced due to these factors, it can misguide the classifier.

The image preprocessing of the dataset was carried out in two specific ways to find the most effective way to emphasize the extractable features of the images. Therefore, several image processing techniques were used in this study to reduce the effect of noises. First, all the images in the dataset are manually categorized into different classes based on the data given.

- Approach 1 – applied a predefined mask – in this image preprocessing method, the original images were cropped to extract the cancer-affected area of the image using a predefined mask set specially designed for the HAM10000 dataset.

Figure 2 illustrates an example mask image designed especially for the AKIEC class. The mask is intended as Portable Network Graphics (PNG) for the easiness of using

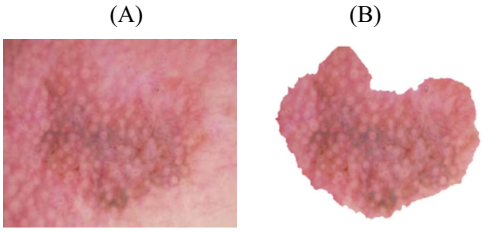
Figure 2
Example image of a mask



the mask over the original image. This allows the user to apply the cover to the image reduction and remove the cancer-affected area of the original image. This step was mainly carried out with the consideration of increasing the model accuracy and performance of the model since the background skin can make a negative impact on the model.

Figure 3 clearly illustrates the before and after the difference between the images after applying the mask over the original image. The main affected area of the image will be cropped and taken out by reducing the unaffected area of the image. Since the image reduction has taken place, the pixel value of the affected area remains as it is in the original image set. As the first approach, the background-reduced images were fed to the feature extraction and the classifier.

Figure 3
(A) Original image and (B) masked image



- Approach 2 – the original image set undergoes approach one as the initial step of this approach, and the masked image set was designed with no disturbance to the cancer-affected area of the image. The output images from approach one feed into a developed filter to improve the image quality by emphasizing the image's cancer-affected location.

  Figure 4 illustrates the procedure followed to improve the quality of the images. Initially, all images were converted from the RGB format to LAB. The converted image was separated into channels L, A, and B. The contrast of channel L was improved with the consideration of the image quality by using the CLAHE (Contrast Limited Adaptive Histogram Equalization). The enhanced images were used to update the original image, and the images were finally converted to the BGR from the LAB.

Figure 4
Filter applying approach

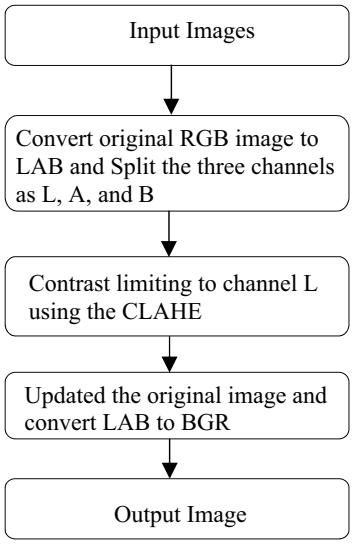
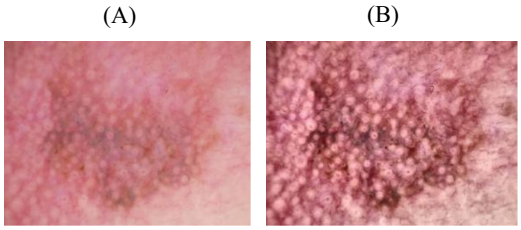


Figure 5 illustrates how the enhanced made up after following the procedure mentioned in Figure 4. Cancer-affected area in the images turned out to be clearer after the filtration process. Applying the enhancing technique to the images made it easier to identify the seven classes with a visual inspection.

The enhanced image was fed to the mask-applying process as shown in Figure 2. After applying the mask to the enhanced, the image was improved, as represented in Figure 5(B).

Figure 5
(A) Original image and (B) enhanced image



Two approaches were considered in this image preprocessing steps to find the best method for skin cancer recognition. Since the images are related to the biomedical side, the best approaches were considered without doing hard manipulations to the dataset.

Figure 6
Explained both approaches

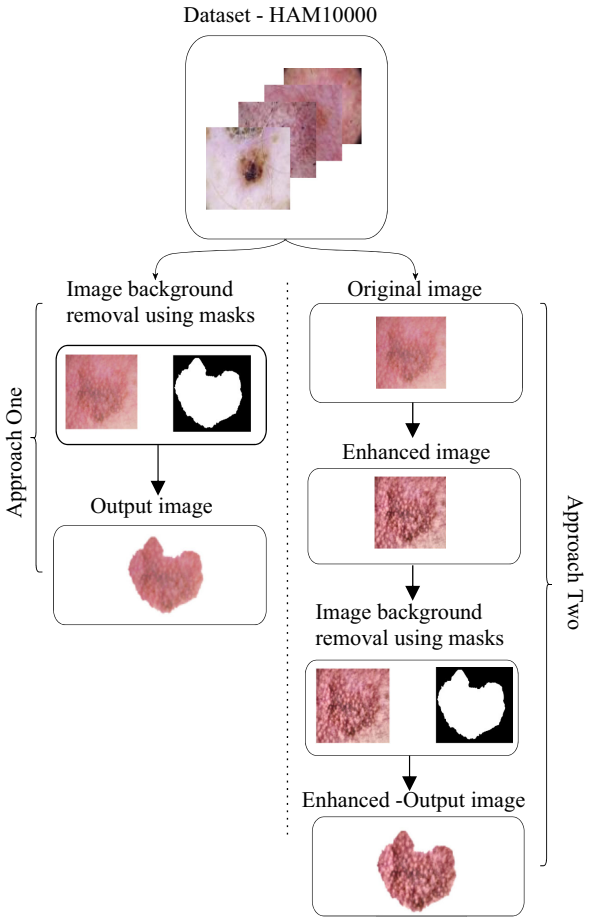


Figure 6 represents both image preprocessing approaches (approach 1 and approach 2) used in this study to identify the best method for the image-based cancer detection system development.

These preprocessed images were then used to identify the interesting points (Mampitiya & Rathnayake, 2022).

The background skin removal process is then applied to all 10015 images available in the dataset, which were gained from both approaches. Through this process, all the available noises in the images were removed and obtained focused areas of each image were for further processing.

3.3. Feature extraction

Feature extraction is a primary step in developing a classifier model as the classifier model learns from the data fed by the feature extraction descriptor.

Edge detection or edge feature extraction in computer vision is used to find the boundaries of an image using pixel location information. In these approaches, it identifies whether a particular pixel is on an edge. This study extracted a HOG features from the images. Those data were passed to the data oversampler. The HOG feature descriptor is used to identify objects in an image, mainly focusing on the structure or the shape of an object. Furthermore, HOG uses magnitude and the gradient angle to compute features; hence, it is identified as a high-performing edge descriptor. According to the image, HOG develops the histogram considering the gradient's magnitude and orientation (Sunguangling, 2008).

As these extracted features are analyzed in localized portions, the precision of the feature extraction is higher. Moreover, the HOG can work accurately and efficiently when the images are focused on a portion or small image.

The HOG was carried out under three main steps (Satpathy et al., 2010; Zhang et al., 2020):

1. Calculation of the gradients – the gradient for the x axis and in the y axis pixels are calculated. The G_x is used for the pixel gradient calculated for the x axis, and G_y are used to demonstrate the pixel gradients calculated for the pixel values in the y axis. $I(x, y)$ are considered as the intensity levels of the pixel location given by the (x, y) coordinates. Therefore, the gradient of the pixel location (x, y) can be gained as:

$$G_x(x, y) = I_{(x+1,y)} - I_{(x-1,y)} \tag{1}$$

$$\mathrm{G_y(x, y)} = \mathrm{I_{(x,y+1)}} - \mathrm{I_{(x,y-1)}} \tag{2}$$

The gradient of the pixel locations can be computed by equation (1) and equation (2). The magnitude of the gradient is represented by the $\mathrm{M(x, y)}$ in equation (3), and the angle $\theta(\mathrm{x, y})$ is represented by equation (4). Then the arctan was calculated to get the orientation:

$$\mathrm{M(x, y)} = \sqrt{\mathrm{G_x(x, y)^2 + G_y(x, y)^2}} \tag{3}$$

$$(\theta) = \arctan\left(\frac{\mathrm{G_{\{y\}}(x,y)}}{\mathrm{G_{\{x\}}(x,y)}}\right) \tag{4}$$

2. Binning of the orientations – a histogram is created by sorting the gradient magnitude of each pixel in a cell into distinct orientation bins based on its gradient angle.
3. Normalization – the histogram creation based on the image gradient is addressed in step 3. The image gradient is dependent on the overall illumination of the image. Nevertheless, ideally, the descriptor should be unaffected by changes in lighting. Hence, the histogram will be unaffected by lightning variations after normalizing.

HOG divides an image into a set of small square cells, generates a histogram of oriented gradients for each cell, normalizes the result using a block-wise sequence, and outputs a descriptor for each cell (Liu, 2021).

Following the steps mentioned above, in the proposed approach, HOG feature descriptor is applied on all 10015 images. The extracted HOG features returned as a matrix with the shape of $(10015,\ 7957)$ was formed.

The matrixes gained by applying the HOG feature extraction to both approaches were equal in size, and the samples were changed. Both the matrixes were in the shape of (10015, 7957). The extracted features from both approaches were sent to the data oversampler.

3.4. Oversampling of the dataset

The HAM1000 is an imbalanced dataset since the largest class NV consists of more than 50% of images from the dataset, and the smallest class DF only consists of as low as 115 images from the dataset. The imbalance of the data under each class negatively impacts the performance of the classifier which will cause to the reduction of the overall performance of the classifier (Flores et al., 2018; Srinilta & Kanharattanachai, 2021). As a solution, oversampling and undersampling of the data can be introduced which will add or remove data from the dataset. This is illustrated in Figure 1.

Synthetic Minority Oversampling Technique (SMOTE) is used in this study to avoid misguidance by the imbalanced dataset by generating synthetic instances of minority classes. The SMOTE approach starts with a small sample of minorities and then creates an equivalent number of synthetic instances (Tallo & Musdholifah, 2018).

Table 1 presents the number of images of each class before and after applying SMOTE. After oversampling the dataset, the data available in each category is balanced to an equal number of 6702. Hence, the dataset shape was enlarged from $(10015,\ 7957)$ to $(46914,\ 7957)$. Since this study focuses on two approaches, the data obtained from the HOG feature extraction for both approaches were sent for the SMOTE to make a balanced dataset. The output matrix gained from the SMOTE for both approaches is the same size $(46914,\ 7957)$.

Table 1
Summary of data after oversampling

Category	Original	After SMOTE
AKIEC	327	6702
BCC	514	6702
NV	6705	6702
BKL	1099	6702
MEL	1113	6702
DF	115	6702
VASC	142	6702
Total	10015	46914

3.5. Feature dimension reduction

The proposed model's performance mainly relies on the data quality used to train the model. Through the SMOTE process, the obtained final shape of the dataset was (46914, 7957). The higher dimensions led to low performance in the model and lesser accuracy. Therefore, the dimension of the oversampled dataset is reduced using principal component analysis (PCA). As explained in the study (Mampitiya et al., 2021), the implementation of the PCA-based dimension reduction process undergoes four main steps:

1. Data standardizing.
2. A covariance matrix was developed by analyzing the features.
3. Eigendecomposition of the covariance matrix.
4. The eigenvectors are sorted by their eigenvalues.

Table 2 elaborates on the dataset dimension's development after applying PCA to the oversampled dataset. After using PCA, the computational complexity is reduced as the dimensions that need to be considered during the computation are lesser.

In this study, the six PCA levels were considered to identify the best PCA level for the dataset. Those levels are 10, 100, 1000, 2000, 4000, and 7955. The dataset matrix changed every time the PCA level changed. This step was carried out to verify and identify the values best fit the ML algorithms.

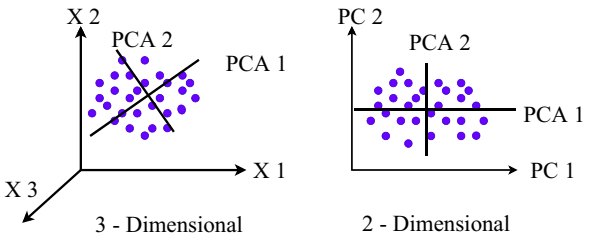
Table 2
Summary of data with PCA technique

PCA (n)	Category	Matrix shape	
		Before PCA	After PCA
10	Training	(32839, 7955)	(32839, 10)
	Testing	(14075, 7955)	(14075, 10)
100	Training	(32839, 7955)	(32839,100)
	Testing	(14075, 7955)	(14075, 100)
1000	Training	(32839, 7955)	(32839, 1000)
	Testing	(14075, 7955)	(14075, 1000)
2000	Training	(32839, 7955)	(32839, 2000)
	Testing	(14075, 7955)	(14075, 2000)
4000	Training	(32839, 7955)	(32839, 4000)
	Testing	(14075, 7955)	(14075, 4000)
7955	Training	(32839, 7955)	(32839, 7955)
	Testing	(14075, 7955)	(14075, 7955)

Table 2 illustrates how the matrix shape changed with the n components used with the PCA. For both approaches, these PCA levels were used, and the accuracies were gained for six PCA levels.

The application and the basic background of undergoing process of the PCA are represented in Figure 7.

Figure 7
Representation of principal component analysis

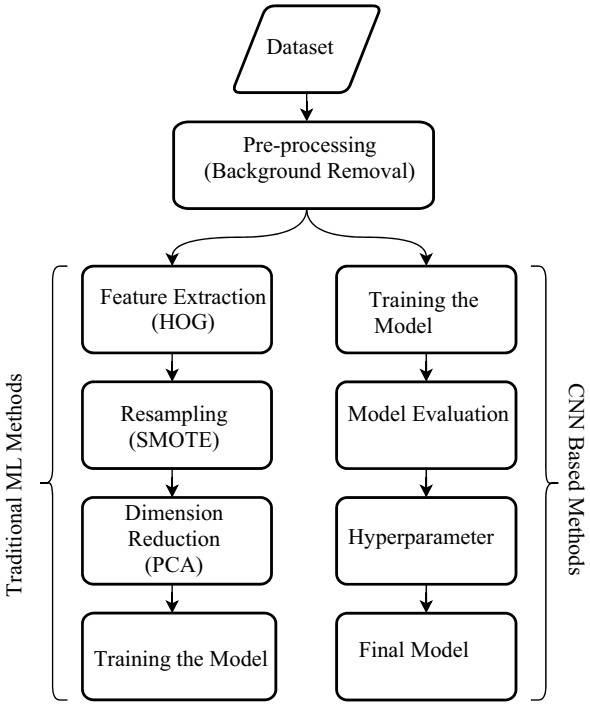


3.6. Implementation of the ML models

As indicated, this study focuses on CNN algorithms and traditional ML algorithms for skin cancer classification. Therefore, the experiments were performed in two ways.

The flow of the experimental procedure is shown in Figure 8. Under the traditional methods, eight well-known ML models were considered. The CNN-based methods that were considered in this study are ResNet50 and MobileNet.

Figure 8
Overall diagram of the evaluation methods



CNN-based ML methods

- *ResNet50:* Residual network is referred to as ResNet. The 2015 computer vision research article "Deep Residual Learning for Image Recognition" introduced this novel neural network. The implementation of the ResNet50 can be found in the study (He et al., 2016). There are other versions of ResNet that use the same basic idea but have varying amounts of layers. The form that can operate with 50 neural network layers is referred known as Resnet50.
- *MobileNet:* MobileNet is a type of CNN that Google open-sourced. The MobileNet model, as its name suggests, is TensorFlow's first mobile computer vision model and is intended for usage in mobile applications. MobileNet uses depth-wise separable convolutions. Compared to a network with ordinary convolutions of the same depth in the nets, it dramatically decreases the number of parameters. Lightweight deep neural networks are the outcome of this method. This feature offers a great place to start when training the classifiers, which are ridiculously tiny and unbelievably quick (Howard et al., 2017) Both the ResNet50 and MobileNet work on the images of the dataset without undergoing the feature maps of the images.

Traditional ML methods

- *Support vector machine (SVM):* SVM represents the supervised machine models and can be used for classification and regression. SVM is mainly based on the conception of the decision planes. Decision planes, also known as hyperplanes, are the border that separates the classes according to the behavior of the data when the input dataset is plotted in an n-dimensional space. The kernel trick is a technique that is used by SVM for the transformation of low-dimensional data into high-dimensional data. The SVM is developed with various kernel tricks, but the most common kernel tricks are linear, polynomial, radial basis function (RBF), and sigmoid (Alam et al., 2016; Mohan et al., 2020):

$$K(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{\|\mathbf{x}-\mathbf{x}'\|^2}{2\sigma^2}\right) \tag{5}$$

Equation (5) is used for the calculation of the RBF kernel for the SVM. The σ is the hyperparameter, and the $K(\mathbf{x} - \mathbf{x}')$ is the Euclidean distance in between the two vectors, and the $K||x, x'||$ is a function that is defined with two input samples x and x' as vectors.

The (γ) value for the SVM was selected as "auto." With that option, the SVM tends to use (1/n_components) of the dataset. So, according to that, the SVM will be able to adjust the (γ) parameter according to the dataset.

Regularization parameters (C) were set to C = 10. The value was identified after setting different values starting from C = 1.

- *K-nearest neighbors (KNN):* KNN is a nonparametric, supervised ML model that can solve classification problems. The KNN algorithm predicts the values of new input data points based on feature similarity. The incoming data point will be assigned a value based on how identical it is to the training dataset, considering the neighbor category. Euclidean distance is the most commonly used distance metric (Chethana, 2021; Huang et al., 2018).
- *Decision tree (DT)*: A decision tree is a learning model that divides an input dataset into subgroups based on an attribute value test. The process will continuously carry on for every derived subset in a recursive manner. The computational power used by the DT is lower than other considered models in this research (Yang, 2019).
- *Naive Bayes (NB):* NB is a supervised classification algorithm based on *Baye's theorem*. The NB is divided into three separate models according to the functioning methods. They are Gaussian Naive Bayes (GNB), Multinomial Naive Bayes (MNB), and Bernoulli. GNB and MNB were used as classifiers in this study (Myaeng et al., 2006).
- *Logistic regression (LR):* This ML regression algorithm is a supervised learning model that predicts a target variable's probability. In a classification of LR, the output variable takes only discrete values for a given set of features (Zou et al., 2019).

All the ML models above were implemented to identify the best model for detecting skin cancers. Feature extraction, oversampling, and feature dimension reduction techniques were introduced to increase the performance of the developing ML models and overcome the traditional problems that ML models face.

We obtained the classical ML models' hyperparameters according to the test data accuracies. The models were trained under different hyperparameters and used the trial-and-error method for the hyperparameter optimization. Using the testing data, the accuracies for the different model conditions were observed, and the best parameters were selected that give out the highest accuracy.

ResNet50 and MobileNet hyperparameters were tuned accordingly, and Table 3 shows the conclusive results of the tuned parameters.

Table 3
Tuned hyperparameters for CNN-based methods

Hyperparameters	MobileNet	ResNet50
Optimizer	Adam	Adam
Learning rate	0.001	0.001
Epochs	100	100
Dropout	0.25	0.40

4. Results and Discussion

This research aimed to develop an accuracy that improved ML combination model to detect skin cancer by using images to input the model. The eight ML models used for the classification of skin cancer images were evaluated using following evaluation parameters:

1. Accuracy (Acc)
2. Precision (Prec)
3. Recall
4. F1-score

Accuracy is a metric that represents the number of data points correctly predicted by the models out of all data points. Precision is a statistic that measures how many of the models produced correct optimistic predictions out of all potential positive predictions. The F1-score calculates the meaning of the precision and the recall. The model will gain a high F1-score value if the precision and the recall values are higher.

Out of all possible positive forecasts, the recall is a metric that counts how many correct optimistic predictions account for the correct positive predictions out of all positive predictions. Recall considers the missed positive predictions instead of precision, which only considers the correct ones out of all positive ones.

The results were categorized according to the PCA level and the algorithms used for the two approaches.

Accuracy, Kappa score, and the Matthew's coefficient were used for this step to verify the performance of the model rather than only depending only on the accuracy of the model. Concerning Table 4, it can be identified that all six algorithms for both approaches perform well when the PCA value is 2000. Furthermore, these results illustrate that the SVM performs well in both approaches at all PCA levels.

The variation of the ML model accuracies with the six PCA levels for both approaches can be identified according to Figure 9. Furthermore, it illustrates that the SVM performs well, beginning from the $n = 1000$ level. Figure 10 illustrates how the accuracies of the ML models will change for the different PCA values for approach 2 data. The next highly accurate model for approach 1 data is LR. Here, the SVM outperforms the remaining ML models by achieving the maximum accuracy of 98.89%. The next highest-performing model after the SVM can be identified as LR.

Considering all PCA ranges with the six different ML models, it can be identified that the SVM is the best-performing algorithm for the extracted HOG features from skin cancer images.

Table 4
Evaluation parameters for ML models under both approaches

		Approach 1						Approach 2					
PCA	Para	SVM	KNN	DT	NBG	NBM	LR	SVM	KNN	DT	NBG	NBM	LR
10	Acc	66.12	71.61	69.4	21.67	17.2	18.05	65.22	71.71	69.3	18.93	15.71	18.06
	Kappa	0.604	0.66	0.64	0.085	0.03	0.04	0.594	0.67	0.64	0.05	0.02	0.01
	Mat	0.609	0.67	0.64	0.08	0.03	0.04	0.598	0.67	0.64	0.08	0.02	0.04
100	Acc	96.16	82.44	72.29	31.42	25.48	30.89	96.31	82.32	72.40	33.04	26.82	31.56
	Kappa	0.955	0.795	0.676	0.199	0.130	0.193	0.956	0.739	0.678	0.218	0.146	0.201
	Mat	0.955	0.805	0.677	0.206	0.133	0.195	0.957	0.804	0.678	0.224	0.15	0.204
1000	Acc	98.69	80.71	68.38	43.07	63.02	78.29	98.62	80.73	67.72	42.90	65.00	78.07
	Kappa	0.986	0.774	0.631	0.335	0.568	0.746	0.984	0.774	0.623	0.333	0.591	0.744
	Mat	0.984	0.791	0.631	0.342	0.570	0.747	0.984	0.790	0.623	0.340	0.594	0.745
2000	Acc	98.85	77.40	71.60	49.40	72.78	87.65	98.89	77.49	72.15	49.1	75.36	88.07
	Kappa	0.986	0.7362	0.668	0.410	0.680	0.855	0.987	0.737	0.675	0.406	0.712	0.860
	Mat	0.986	0.764	0.669	0.413	0.680	0.857	0.987	0.765	0.675	0.409	0.715	0.862
4000	Acc	98.79	75.73	75.91	58.2	81.2	92.05	98.66	76.56	76.9	55.42	83.71	92.45
	Kappa	0.987	0.7167	0.719	0.512	0.780	0.907	0.984	0.726	0.730	0.48	0.81	0.911
	Mat	0.986	0.753	0.719	0.514	0.782	0.910	0.984	0.756	0.730	0.482	0.812	0.914
7955	Acc	98.11	87.47	72.75	55.44	86.33	90.35	97.99	88.07	75.57	55.49	88.88	90.62
	Kappa	0.977	0.853	0.682	0.480	0.840	0.887	0.976	0.860	0.715	0.480	0.870	0.89
	Mat	0.978	0.862	0.684	0.483	0.845	0.893	0.976	0.869	0.715	0.483	0.875	0.896

Figure 9
ML model accuracies for approach 1 with PCA levels

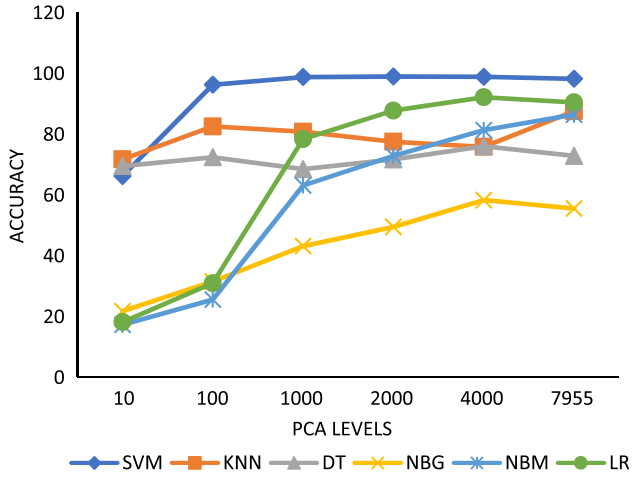
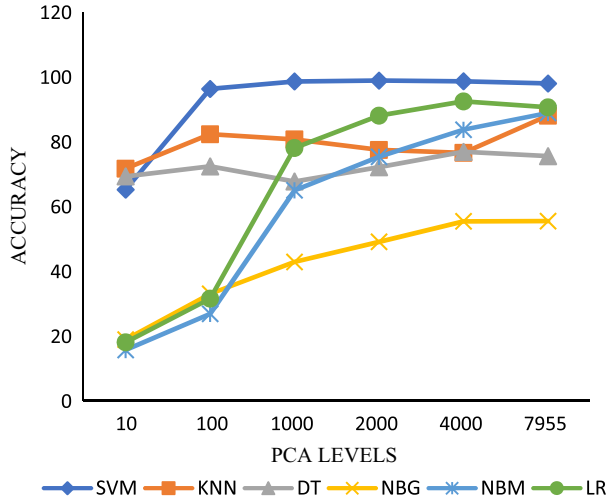


Figure 10
ML model accuracies for approach 2 with PCA levels



Moreover, considering Figures 9, 10, and Table 4, it can be identified that 2000 is the best value suitable for the PCA.

With the comparison of Table 4, approaches 1 and 2, the results emphasize that approaches 1 and 2 give nearly the same results for all six ML models, while SVM performs well in both approaches. The accuracies for approach 1 SVM are 98.85%, while approach 2 SVM is 98.89%.

Moreover, K-fold cross-validation was conducted on SVM to ensure the results. The validation accuracy is achieved with the validation data (Tong et al., 2016). The statistical graph shown in Figure 11 represents the accuracy variation with the increment of the K value. The K-fold cross-validation was carried out up to 10 K values. Figure 11 further illustrates that at K = 2, approach 1 SVM model attained the highest accuracy of 99.15% and achieved the lowest accuracy of 98.55% at K = 6. For approach 2 SVM, the highest accuracy was gained as 99.05% at the initial K value. The lowest accuracy was gained as 98.03% at K = 6. Considering Figure 11, K-fold cross-validation values for the SVM algorithm can be verified that approach 1 performs better than approach 2, where the filter was applied, as illustrated in Figure 6.

Figure 11
K-fold value for SVM (approach 1 and 2)

Table 5
Comparison summary of results for all machine learning models (PCA = 2000, approach 1)

ML model		ACC	Prec	Recall	F1
CNN-based	ResNet50	83.00	0.81	0.83	0.78
	MobileNet	72.00	0.86	0.72	0.77
Traditional	SVM	99.15	0.99	0.99	0.99
	KNN	77.40	0.79	0.83	0.77
	DT	71.60	0.71	0.71	0.71
	G-NB	49.4	0.49	0.49	0.49
	M-NB	72.78	0.72	0.72	0.72
	LR	87.65	0.87	0.88	0.87

The accuracy of the SVM model for all seven available classes was evaluated via the confusion matrix with the consideration of class-level accuracies. The confusion matrix is a graphical method that represents the accuracies gained by each class by considering the actual class and the model-predicted class. With the analysis of the evaluation as mentioned above parameters in Table 5 that tabulated for approach 1 data, it can be identified that a combination of skin background removal, HOG feature extraction, SMOTE, and PCA with the SVM outperformed the rest of the algorithms used in this study.

This evaluating graphical method represents how the model is well trained to identify the classes of the dataset. Figure 12 represents the class-wise accuracies attained by the proposed method used with the SVM.

Table 6 represents that three out of seven classes attained a maximum accuracy of 100% and the class NV achieved a minimum accuracy of 97.75%.

The computational cost for the CNN and the SVM models was considered about the consuming time for the training and the amount of resources used for the computation. The SVM used around 2 hours and 23 minutes to train the model. For the training of the two CNN models, the time was more than 3 hours.

Regarding resource usage, this research study uses 24 GB RAM and Tesla K80. The CNN models used more resources than the SVM model considering the computation resources.

Figure 12
Confusion matrix for the SVM

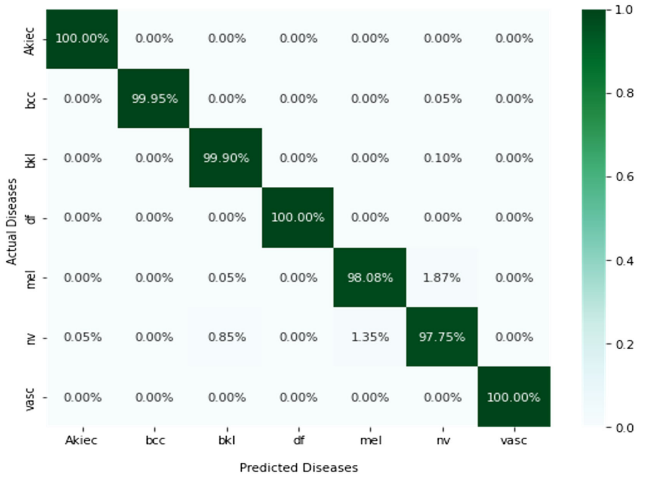


Table 6
Class-wise accuracies attained by SVM

Class	Accuracy (%)
AKIEC	100
BCC	99.95
BKL	99.90
DF	100
MEL	98.08
NV	97.75
VASC	100

Table 7
Comparison of accuracy with similar works

Reference model	Algorithm	Accuracy
This study	*SVM*	*99.15%*
Moldovan (2019)	Transfer learning	85.00%
Shah (2021)	LRNet	90.60%
Huo (2021)	CNN	75.00%
Waweru et al. (2020)	DCNN	78.00%
Cakmak and Tenekeci (2021)	Nasnet Mobile	97.90%

5. Conclusion and Future Work

This study seeks to fill the research gap by designing and implementing low-cost and lower computational powered skin cancer classification systems with higher accuracy. The discussed methodology uses two approaches with several steps to achieve comparative accuracy compared with the related studies, such as applying mask, filtering, HOG feature extraction, PCA dimension reduction, SMOTE oversampling technique, and model training.

According to the results, it can be identified that the SVM model is efficient and works at a higher accuracy for both approaches at every PCA level. At the same time, the best performance was gained at PCA level 2000. Moreover, with the cross-validation, it was confirmed that approach 1 with SVM performs well than approach 2. The accuracy of the SVM model outperformed well-known CNN-based image classification models such as ResNet50 and MobileNet. Moreover, Table 7 summarizes the past studies of skin cancer classification using various ML algorithms. It can be identified that the SVM method for image-based skin cancer detection performs well compared with the other models mentioned, with higher accuracy of 99.15% by classifying seven well-known skin cancers. The developed model can identify the above-mentioned seven skin cancers that arise in the external skin of the human body. Future studies in this field might include developing a real-time visual-based skin cancer detection system using a proposed combination for SVM.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

References

Alam, S., Kang, M., Pyun, J. Y., & Kwon, G. R. (2016). Performance of classification based on PCA, linear SVM, and Multi-kernel SVM. In *International Conference on Ubiquitous and Future Networks*, 987–989. https://doi.org/10.1109/ICUFN.2016.7536945

Cakmak, M., & Tenekeci, M. E. (2021). Melanoma detection from dermoscopy images using Nasnet Mobile with transfer learning. In *2021 29th Signal Processing and Communications Applications Conference,* 1–4. https://doi.org/10.1109/SIU53274.2021.9477985

Chethana, C. (2021). Prediction of heart disease using different KNN classifier. In *Proceedings – 5th International Conference on Intelligent Computing and Control Systems*, 1186–1194. https://doi.org/10.1109/ICICCS51141.2021.9432178

Demir, A., Yilmaz, F., & Kose, O. (2019). Early detection of skin cancer using deep learning architectures: Resnet-101 and inception-v3. In *2019 Medical Technologies Congress*, 1–4. https://doi.org/10.1109/TIPTEKNO47231.2019.8972045

Flores, A. C., Icoy, R. I., Pena, C. F., & Gorro, K. D. (2018). An evaluation of SVM and naive bayes with SMOTE on sentiment analysis data set. In *ICEAST 2018 – 4th International Conference on Engineering, Applied Sciences and Technology: Exploring Innovative Solutions for Smart Society*, 1–4 https://doi.org/10.1109/ICEAST.2018.8434401

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778.

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., ... & Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv Preprint: 1704.04861*.

Huang, J., Wei, Y., Yi, J., & Liu, M. (2018). An improved knn based on class contribution and feature weighting. *Proceedings – 10th International Conference on Measuring Technology and Mechatronics Automation*, 313–316. https://doi.org/10.1109/ICMTMA.2018.00083

Huo, Y. (2021). Full-stack application of skin cancer diagnosis based on CNN model. In *2021 IEEE International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology*, 754–758. https://doi.org/10.1109/CEI52496.2021.9574583

Liu, Z. (2021). Improvement of feature extraction based on HOG. In *Proceedings – 2021 2nd Asia Symposium on Signal Processing*, 55–60. https://doi.org/10.1109/ASSP54407.2021.00017

Mampitiya, L. I., Nalmi, R., & Rathnayake, N. (2021). Classification of human emotions using ensemble classifier by analysing

EEG signals. In *Proceedings – 2021 IEEE 3rd International Conference on Cognitive Machine Intelligence*, 71–77. https://doi.org/10.1109/COGMI52975.2021.00018

Mampitiya, L. I., & Rathnayake, N. (2022). An efficient ocular disease recognition system implementation using GLCM and LBP based multilayer perception algorithm. In *2022 IEEE 21st Mediterranean Electrotechnical Conference*, 978–983. https://doi.org/10.1109/MELECON53508.2022.9843023

Mohan, L., Pant, J., Suyal, P., & Kumar, A. (2020). Support vector machine accuracy improvement with classification. In *Proceedings – 2020 12th International Conference on Computational Intelligence and Communication Networks*, 477–481. https://doi.org/10.1109/CICN49253.2020.9242572

Moldovan, D. (2019). Transfer learning based method for two-step skin cancer images classification. In *2019 E-Health and Bioengineering Conference*, 1–4. https://doi.org/10.1109/EHB47216.2019.8970067

Myaeng, S. H., Han, K. S., & Rim, H. C. (2006). Some effective techniques for naive bayes text classification. *IEEE Transactions on Knowledge and Data Engineering*, *18*(11), 1457–1466. https://doi.org/10.1109/TKDE.2006.180

Satpathy, A., Jiang, X., & Eng, H. L. (2010). Extended histogram of gradients feature for human detection. In *Proceedings – International Conference on Image Processing*, 3473–3476. https://doi.org/10.1109/ICIP.2010.5650070

Shah, M. (2021). LRNet: Skin cancer classification using low-resolution images. In *Proceedings – International Conference on Communication, Information and Computing Technology*, 1–5. https://doi.org/10.1109/ICCICT50803.2021.9510138

Srinilta, C., & Kanharattanachai, S. (2021). Application of natural neighbor-based algorithm on oversampling SMOTE algorithms. In *2021 7th International Conference on Engineering, Applied Sciences and Technology*, 217–220. https://doi.org/10.1109/ICEAST52143.2021.9426310

Subha, S., Wise, D. C. J. W., Srinivasan, S., Preetham, M., & Soundarlingam, B. (2020). Detection and differentiation of skin cancer from rashes. In *Proceedings of the International Conference on Electronics and Sustainable Communication Systems*, 389–393. https://doi.org/10.1109/ICESC48915.2020.9155587

Sunguangling, Q. (2008). Handwritten character recognition using multiresolution histograms of oriented gradients. In *International Conference on Communication Technology Proceedings*, 715–717. https://doi.org/10.1109/ICCT.2008.4716216

Tallo, T. E., & Musdholifah, A. (2018). The implementation of genetic algorithm in smote (synthetic minority oversampling technique) for handling imbalanced dataset problem. In *Proceedings – 2018 4th International Conference on Science and Technology*. 1-4. https://doi.org/10.1109/ICSTC.2018.8528591

Tong, J., Guo, Q., Xi, J., Yu, Y., & Schreier, P. J. (2016). Choosing the diagonal loading factor for linear signal estimation using cross validation. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing*, 3956–3959. https://doi.org/10.1109/ICASSP.2016.7472419

Tschandl, P., Rosendahl, C., & Kittler, H. (2018). Data descriptor: The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific Data*, *5*(1), 1–9. https://doi.org/10.1038/SDATA.2018.161

Waweru, A. K., Ahmed, K., Miao, Y., & Kawan, P. (2020). Deep learning in skin lesion analysis towards cancer detection. In *Proceedings of the International Conference on Information Visualisation*, 740–745. https://doi.org/10.1109/IV51561.2020.00130

Yang, F. J. (2019). An extended idea about decision trees. In *Proceedings – 6th Annual Conference on Computational Science and Computational Intelligence*, 349–354. https://doi.org/10.1109/CSCI49370.2019.00068

Zhang, L., Zhou, W., Li, J., Li, J., & Lou, X. (2020). Histogram of oriented gradients feature extraction without normalization. *Proceedings of 2020 IEEE Asia Pacific Conference on Circuits and Systems*, 252–255. https://doi.org/10.1109/APCCAS50809.2020.9301715

Zou, X., Hu, Y., Tian, Z., & Shen, K. (2019). Logistic regression model optimization and case analysis. *Proceedings of IEEE 7th International Conference on Computer Science and Network Technology*, 135–139. https://doi.org/10.1109/ICCSNT47585.2019.8962457

World Health Organization. (2022). Questions and answers. Retrieved from https://www.who.int/news-room/questions-and-answers/

How to Cite: Mampitiya, L. I., Rathnayake, N., & De Silva, S. (2023). Efficient and Low-Cost Skin Cancer Detection System Implementation with a Comparative Study Between Traditional and CNN-Based Models. *Journal of Computational and Cognitive Engineering* 2(3), 226–235, https://doi.org/10.47852/bonviewJCCE2202482