STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **12**, November 2024, pp 1921–1933.
Published online in International Academic Press (www.IAPress.org)

Skin Cancer Diagnosis With Multi-Level Classification

Rania El-Badawy [1], BenBella S. Tawfik [2], Mohamed Amal Zeidan [1,*]

[1] *Engineering College- Zagazeg University*
[2] *Faculty of Computer and Informatics - Suez Canal University*

Abstract Skin cancer arises from the uncontrolled proliferation of abnormal skin cells, primarily triggered by exposure to the harmful ultraviolet (UV) rays of the sun and the utilization of UV tanning beds. This condition poses a heightened risk due to its potential to progress into blood cancer and lead to rapid fatality. Extensive research efforts have been dedicated to advance the treatment of this perilous disease. This paper presents a system designed for the examination and diagnosis of pigmented skin lesions and melanoma.
The system incorporates a supervised classification algorithm that combines Convolutional Neural Network (CNN) and Deep Neural Network (DNN) architectures with feature extraction techniques. It operates in two distinct stages: the initial stage classifies images into two categories, namely benign or malignant, while the subsequent stage further categorizes the images into one of three classes: basal cell carcinomas, squamous cell carcinomas, or melanoma. Consequently, the comprehensive system addresses four classes, namely benign, basal cell carcinomas, squamous cell carcinomas, and melanoma.
This work contributes to the system's design in three significant ways. Firstly, it implements multiple iterations to select the most optimal images, resulting in the highest classification accuracy. Secondly, it employs various statistical methods to identify the most pertinent features, thereby enhancing the classifier's accuracy by focusing on the most informative features for the classification task. Lastly, a two-stage classification approach is implemented, employing two distinct classifiers at different levels within the overall system. Despite the inherent complexity of the real-world problem, the overall system attains a commendable level of classification accuracy.
Following rigorous experimentation, the study identifies the top three models. Each approach culminates in a classifier for each stage. The first approach, utilizing a deep learning classifier, achieves an accuracy of 86.08% in the initial cancer discrimination stage and 79.12% in the subsequent stage. The second approach, employing a machine learning classifier, attains an accuracy of 74.63% in the first stage and 64.41% in the second stage. The third approach, utilizing a linear regression classifier, achieves an accuracy of 98% in the first stage and 90% in the second stage. These results underscore the significance of feature selection in influencing model accuracy and suggest the potential for further optimization.

Keywords skin cancer, melanoma, deep learning, neural network, feature extraction, dermoscopy images.

DOI: 10.19139/soic-2310-5070-2090

1. Introduction

There are three primary categories of skin cancer: Basal Cell Carcinomas (BCCs), squamous cell carcinomas (SCCs), and melanoma [3].
BCCs represent the initial stage of skin cancer and occur when UV radiation-induced DNA damage triggers alterations in basal cells, leading to uncontrolled cell growth. These cancers often manifest as open sores, red patches, pink growths, shiny bumps, scars, or growths with raised edges and a central depression. Typically, BCCs appear on sun-exposed areas of the body, such as face, and may exhibit pigmentation in individuals with darker skin.

*Correspondence to: Mohamed Amal Zeidan (Email: Dr.mohamedamalzeidan@gmail.com), Engineering College - Zagazeg University

ISSN 2310-5070 (online) ISSN 2311-004X (print)
Copyright © 2024 International Academic Press

SCCs constitute the second stage of skin cancer, originating from flat cells in the top layer of the skin. They are commonly found on sun-exposed regions like the head, neck, chest, upper back, ears, lips, arms, legs, and hands. Unlike other skin cancer types, SCCs have the potential to spread to nearby tissues, bones, and lymph nodes, rendering treatment more challenging when not diagnosed early.

Melanoma, the third stage of skin cancer, poses a greater threat when not detected promptly [12]. It arises from damaged melanocytes, often resembling moles and appearing on any part of the body, even areas not exposed to the sun. Melanoma is more prevalent among individuals with fair skin and is the most lethal form of skin cancer. Dermoscopy, a noninvasive high-resolution imaging technique, aids physicians in achieving more accurate skin cancer diagnoses. Due to its aggressive growth, malignant melanoma ranks as one of the fastest-spreading cancers globally. Once it metastasizes to other tissues, the response rate to treatment can plummet to as low as 5%, and the 10-year survival rate hovers at only about 10%. Unfortunately, surgical removal is no longer an option after metastasis occurs, underscoring the critical importance of early detection. Among various skin cancers, melanoma exhibits the highest false negative rate.

In conclusion, precise skin cancer type diagnosis is crucial for determining the appropriate treatment approach. While most research focuses on distinguishing benign from malignant skin lesions, this study aims to classify lesions into two stages. In the first stage, lesions are categorized as benign or malignant, and if malignant, their specific type is identified in the second stage. Various approaches for skin cancer detection have been introduced, but the complexity of the input data images poses significant challenges. This research presents the best-performing models in this context.

This paper is organized as follows: Section 2 provides a brief review of the related work. Summary of the problem definition and contributions is provided in section 3. Section 4 introduces the selected best models. Section 5 offers a comprehensive description of the input database. Section 6 outlines the implementation of three different modeling approaches using this database, and the analysis of the results are presented in Section 7. Section 8 discuss the conclusions. Finally, the used references in section 9.

2. Related work

Various research groups have attempted to classify melanoma using different datasets. Most previous studies focused on diagnosing skin lesions into only two categories: benign and malignant, with a maximum achieved accuracy of 95%. However, in recent years, only a limited number of high-quality research works have focused on classifying the specific types of skin cancer. The model in [10] utilizes a Convolutional Neural Network (CNN) to achieve a high accuracy of 98.07% in recognizing skin diseases like Malignant Melanoma, Squamous Cell Carcinoma, and Basal Cell Carcinoma. The proposed DeepCNN network aims to detect epidermis cancer at early stages, featuring a classifier for three types of skin malignancy: Basal Cell Carcinoma (BCC), Squamous Cell Carcinoma (SCC), and Melanoma. The network can provide precise results, including the identification of the infected skin area. In [13], a novel approach focuses on classifying and segmenting skin lesions to develop a skin cancer detection system with minimal error. The method involves using a standard digital camera for high-resolution imaging of skin lesions. The combination of analytical and segmentation methods aims to create an effective interface to assist dermatologists in diagnosis. The study utilizes back-propagated artificial neural networks for the detection and classification of skin cancer, achieving a 92% classification accuracy in experimental results, demonstrating its viability. In [14], the research emphasizes preprocessing, specifically hair removal, to enhance the classification of skin cancers in dermoscopy images. Existing methods have limitations, prompting the use of a multi-resolution joint Gaussian filter and a hybrid decision-based window mask generation approach in the proposed system. This combination effectively denoises dermoscopy images, identifies unwanted regions through self-adaptive thresholding, removes unwanted objects using connected component labeling, determines hair threshold levels with morphological area opening, and employs a fast marching in-painting procedure to eliminate hair and restore the skin. In [5], an automated medical decision support system for skin cancer developed with normal and abnormal classes. First the discrete wavelet transformation was applied on the images to get the

vector feature. Afterwards, those vectors were used for classification either with feed-forward neural network or k-nearest neighbor algorithm. The results of the deployed techniques were promising as, they got 97.5% for accuracy. The models we propose in this study utilize various techniques during the preprocessing stage. In these models, we employ feature selection techniques, where specific features in the dataset, such as gender, age, and the location of the skin disease, are excluded. The exclusion of these features results in a reduction in the accuracy of our proposed models. To elaborate, the data is segregated based on gender and the location of the skin disease, leading to a decrease in the number of samples. Our introduced models strive to use an optimistic minimum number of relevant features and a minimal number of layers in the neural network. In our manuscript, we classified the type of skin lesion as benign or malignant in the first stage of classification, and the accuracy rate reached 98%. If the result is positive for skin cancer, the cancer is then diagnosed at any of the three stages of cancer. The accuracy of diagnosis at this stage reached 90%. Notice that in the proposed models, we used only the testing set to calculate the accuracy of the model (20% of the dataset).

3. Problem definition

The issue addressed in this research is achieving a high accuracy of classification in complex real-world problems using limited knowledge. Typically, additional information such as blood analysis is required to support the physician's correct diagnosis. This research aims to extract the most information possible from the input images.

4. Classification models

In this section the selected best models are introduced and analyzed.

4.1. Deep learning and machine learning

Artificial Intelligence (AI) is a rapidly growing field that involves the creation of computer programs designed to perform tasks that would typically require human intelligence. It involves various sub-fields such as machine learning, neural networks, and deep learning. The latter, being a subfield of machine learning, is based on the concept of artificial neural networks (ANNs) that simulate human brain functions [9]. ANNs are made up of four main components: inputs, weights, a bias, and an output. They function by utilizing algorithms that compute the input data, adjusting the weights and biases accordingly, and producing an output. This process is similar to linear regression, with a formula that relates inputs, weights, and biases to produce an output.

$$\sum_{\mathrm{i}=1}^{m} \omega_i \chi_i = \omega_1 \chi_1 + \omega_2 \chi_2 + \omega_3 \chi_3 + bias \tag{1}$$

Deep learning is a subfield of machine learning that enables computers to learn from experience, much like humans. It is used in various AI applications and services, from image processing to language recognition. The term "deep" in deep learning refers to the depth of layers in a neural network, with most deep neural networks having more than three layers (input, hidden, and output). Unlike traditional machine learning, which requires manual feature extraction from structured and labeled data, deep learning algorithms can process and extract features from unstructured data, such as images and text, without human intervention. This makes it a powerful tool for automation and data analysis.

Deep learning algorithms use gradient descent and backpropagation to fit the model for accuracy, allowing it to make predictions with increased precision. Deep learning models can learn through supervised learning, unsupervised learning, or reinforcement learning. Supervised learning requires labeled data and human intervention to label input data, while unsupervised learning detects patterns in the data without labels and clusters the data based on distinct characteristics. Reinforcement learning, on the other hand, allows the model to learn through feedback in an environment in order to maximize rewards.

In conclusion, deep learning differentiates itself from classical machine learning by its ability to process unstructured data and automatically extract features, removing some dependency on human experts. It also offers different types of learning approaches, making it a versatile tool for solving complex problems in automation and data analysis.

4.1.1. ***Deep learning:*** Deep neural networks are composed of multiple layers of interconnected nodes, where each layer uses the output of the previous layer to refine and optimize its predictions or categorizations [8]. This forward progression of computation is referred to as ”forward propagation” as shown in figure (1). The input and output layers in a deep neural network are referred to as the ”visible layers.” To further improve the accuracy of predictions, another algorithm, known as ”backpropagation,” is utilized in conjunction with gradient descent. Backpropagation calculates errors in the predictions and then adjusts the weights and biases of the network by moving backwards through the layers. This training process allows the neural network to make increasingly accurate predictions by continuously correcting any errors.

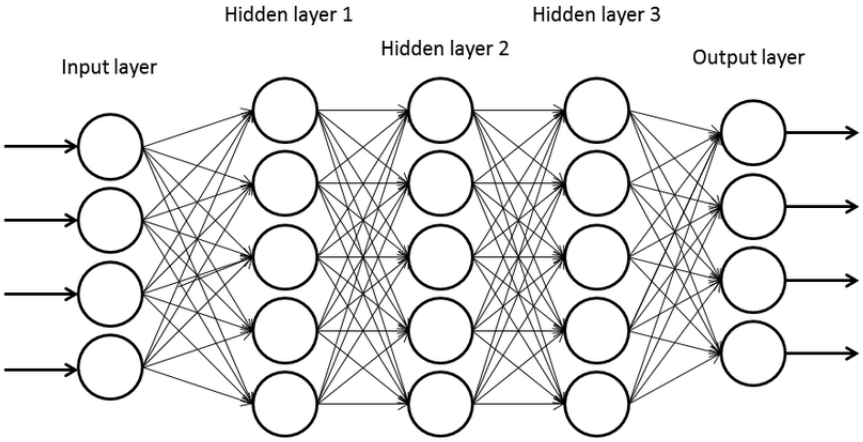

Figure 1. **Deep learning model.**

Convolutional Neural Networks (CNNs) are specialized neural networks designed for computer vision and image classification applications. They are known for their superior performance in image and audio signal processing. A CNN has three main types of layers: convolutional layer, pooling layer, and fully-connected (FC) layer.The convolutional layer is the core building block of a CNN, where the majority of the computation takes place. It has three components: input data, a filter, and a feature map. The input data is a three-dimensional (height, width, depth) matrix representing RGB values in an image. The filter, also known as a kernel or a feature detector, is a two-dimensional array of weights applied to receptive fields in the image. A convolution is performed between the input pixels and the filter, resulting in a dot product that is fed into an output array. The filter then shifts by a stride and repeats the process until the entire image has been swept. The final output is known as the feature map, activation map, or convolved feature. The weights in the feature detector are fixed as it moves across the image, known as parameter sharing. However, some parameters, like the weight values, are adjusted during training through backpropagation and gradient descent. There are also three hyper parameters that affect the volume size of the output, which must be set prior to the training of the neural network: stride, padding, and filter size. The depth of the output can be controlled by the number of filters applied. For instance, if three different filters are applied, then three separate feature maps would be generated, resulting in a depth of three. Stride refers to the number of pixels that the kernel moves over the input matrix. Although stride values of two or higher are not common, a larger stride results in a smaller output. Zero-padding is utilized when the filters do not match the size of the input image. This is done by setting all elements outside of the input matrix

to zero, producing either a larger or an equal-sized output. There are three types of zero-padding that can be applied.

•**Valid padding:** This is also known as no padding. In this case, the last convolution is dropped if dimensions do not align.
•**Same padding:** This padding ensures that the output layer has the same size as the input layer
•**Full padding:** This type of padding increases the size of the output by adding zeros to the border of the input.

After the convolution operation, a CNN applies a Rectified Linear Unit (ReLU) function to the feature map, which adds nonlinearity to the model. The initial convolution layer can be followed by another convolution layer, thus increasing the hierarchical structure of the CNN. Later layers have a wider receptive field, allowing them to process more pixels from the input. Pooling layers, also known as down-sampling, reduce the dimensionality of the input and the number of parameters. In pooling, the kernel applies an aggregation function, such as maximum or average, to the values within the receptive field and stores the result in the output array. Unlike the convolutional layer, the pooling layer does not have any weights. There are two main types of pooling: max pooling and average pooling.

•**Max pooling:** As the filter moves across the input, it selects the pixel with the maximum value to send to the output array. As an aside, this approach tends to be used more often compared to average pooling.
•**Average pooling:** As the filter moves across the input, it calculates the average value within the receptive field to send to the output array.

While a lot of information is lost in the pooling layer, it also has a number of benefits to the CNN. They help to reduce complexity, improve efficiency, and limit risk of overfitting. The final layer in a Convolutional Neural Network (CNN) is the fully-connected layer, as the name implies, each node in the output layer is directly connected to a node in the previous layer. This layer performs the task of classification based on the features extracted through the previous layers and their respective filters. Unlike the convolutional and pooling layers, which usually utilize Rectified Linear Unit (ReLU) activation functions, the fully-connected layer typically employs a softmax activation function to classify inputs, resulting in a probability output ranging from 0 to 1.

4.1.2. ***Machine learning:*** Machine learning is a subfield of AI and computer science that utilizes data and algorithms to imitate human learning. It works in three steps: decision-making, error function evaluation, and model optimization. The algorithm will repeat this process until it reaches a satisfactory accuracy level. There are three main categories of machine learning models: supervised learning, unsupervised learning, and semi-supervised learning. Supervised learning uses labeled datasets to train algorithms to classify and predict outcomes. Some popular techniques include neural networks, Naive Bayes, linear regression, logistic regression, random forests, and SVM. Unsupervised learning analyzes and categorizes unlabeled datasets to find patterns and groupings without human guidance. Algorithms like k-means clustering and neural networks are used in unsupervised learning. Semi-supervised learning combines both supervised and unsupervised learning with a limited set of labeled data and a larger set of unlabeled data. Neural networks and clustering methods are commonly used in semi-supervised learning. In addition to these techniques, other machine learning methods include linear regression, logistic regression, decision trees, random forests, and Support Vector Machines. These algorithms can be used for various real-world applications like predicting numerical values, classifying data into categories, identifying patterns, and more.

4.2. Logistic regression

Logistic Regression is a statistical and machine learning technique used for classification purposes [15]. It determines the relationship between one or more independent variables and a dependent variable, with the goal of finding the best fitting model. For example, when trying to classify whether a certain type of cancer is malignant or not, linear regression cannot be used as it returns continuous values and requires a threshold to convert the

predicted value into a class label. This can lead to incorrect classifications, especially when the data is imbalanced. Linear regression is better suited for regression problems, while logistic regression is specifically designed for classification problems. In logistic regression, the predicted probability is transformed into a class label through the use of a threshold. A value above the threshold indicates one class, and a value below it indicates the other. Logistic regression also addresses the sensitivity to imbalanced data seen in linear regression by using a logistic function to model the relationship between the dependent and independent variables. Logistic regression is widely used in a variety of applications, such as credit scoring, medical diagnosis, and image classification. The logistic regression algorithm uses an optimization process, such as maximum likelihood estimation, to determine the parameters of the model that best fit the data.

4.2.1. ***Linear regression and logistic regression*** This paper presents three classification models for data analysis. Linear regression models examine the correlation between a continuous dependent variable and one or more independent variables. Simple linear regression is used when there is only one independent and dependent variable, while multiple linear regression is used with more than one independent variable. Both linear and logistic regression predict the relationship between dependent and independent variables, with logistic regression specifically designed for predicting categorical outcomes. Linear regression is simple to understand, produces a line of best fit, and uses the least squares method, while logistic regression predicts a probability based on a larger, more representative sample. However, logistic regression requires a larger sample size to detect significant results with sufficient statistical power. Linear regression assumes a linear relationship between the dependent and independent variables, while logistic regression models the relationship as a non-linear transformation of the independent variables. Linear regression is also sensitive to outliers, which can have a significant impact on the model's predictions.

5. Database:

The data for this study was obtained from the International Skin Imaging Collaboration (ISIC Archive). The overall data set is 1200 images. The images were divided into two sets: one set was used for training the models, while the other set was used to evaluate the models' performance. The method of dermoscopy was utilized to diagnose melanocytic lesions, such as nevus and melanoma, by examining various pigmented structures like pigment networks, dots, globules, and streaks. Dermoscopic images provide higher contrast between the lesion and the surrounding skin as shown in figure (2), eliminating the need for additional image enhancement filters, and can greatly reduce the number of unnecessary biopsies. Furthermore, dermoscopy is a non-invasive diagnostic method that allows for a more accurate and efficient diagnosis of skin lesions. Additionally, the use of dermoscopy in this study allowed for a larger, more diverse sample size, which enhances the validity and reliability of the results.

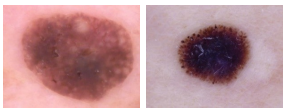

Figure 2. (a) and (b) are dermoscopic images (a is a benign and b is cancer).

6. The models:

This paper will provide an in-depth examination of three models that operate in two stages. The first stage involves dividing the input data into two classes, benign and malignant. The second stage involves further classifying the malignant class into three types of skin cancer.

6.1. First Approach-deep learning classifier:

This model utilizes supervised deep learning to classify the input data in two steps: extracting relevant features from the input images automatically and then classifying them. The input data is divided into 80% for training and 20% for testing. The following steps are involved in the algorithm:

following steps are involved in the algorithm:

1. Loading the data set and storing it in a 3D array.
2. Converting the input to grayscale images.
3. Resizing all images to be of the same size (128×128).
4. Dividing the data into training (80%) and testing (20%) sets.
5. Defining the model architecture, including convolution, max normalization, and ReLU layers, all of which are fully connected.
6. Training the network.
7. Calculating the accuracy of the model

$$accuracy = (numperofcorectedpredection)/(totalnumberofpredection) * 100 \quad (2)$$

The accuracy of the model in the first stage of 86.08%, as shown in figure (3). And the accuracy of the second stage is 79.12%, as shown in figure (4). From this results we find that this algorithm is suitable in classification between two classes but the result in the second stage is very bad.

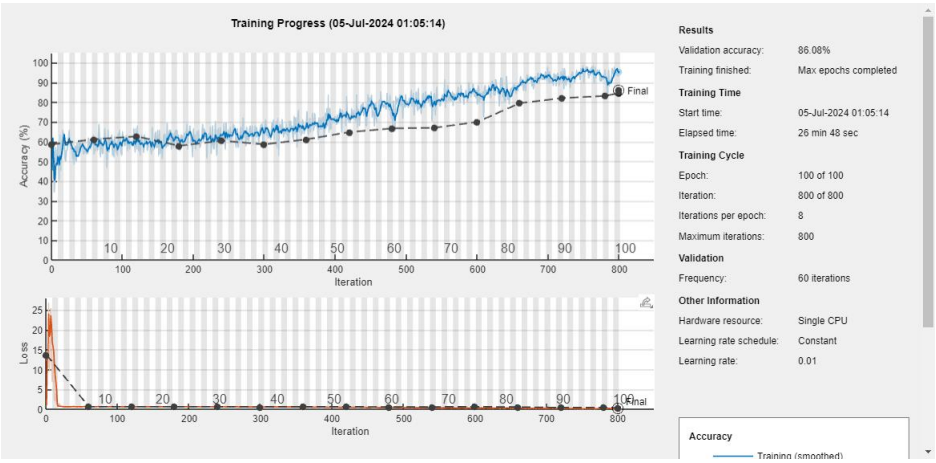

Figure 3. The output training progress for classification between two classes (benign or malignant.)

6.2. Second Approach-machine learning classifier:

Feature extraction involves reducing the number of resources required to describe a large set of data. When performing analysis of complex data one of the major problems stems from the number of variables involved. Analysis with a large number of variables generally requires a large amount of memory and computation power, also it may cause a classification algorithm to over fit to training samples and generalize poorly to new samples. A popular approach to this problem of high-dimensional datasets is to search for a projection of the data onto a smaller number of variables (or features) which preserves the information as much as possible. This model employs a machine learning approach for image classification. Instead of using raw images as inputs, the model inputs features extracted from the images. The first set of features used are the Gray-level Co-occurrence Matrix (GLCM) features. The GLCM is a statistical method of evaluating image texture that takes into account the spatial relationship of pixels [2]. By calculating the frequency of occurrence of pairs of pixels with specific values and in a specific spatial relationship, a GLCM matrix is created. Then, statistical measures are derived from this matrix, including contrast, correlation, energy, and homogeneity.

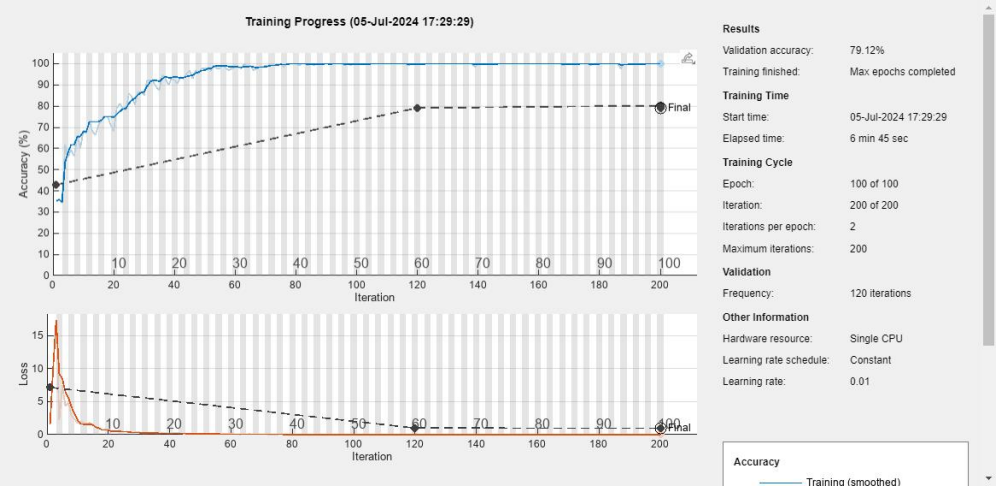

Figure 4. The output training progress for classification between three skin cancer classes.)

Contrast, in particular, measures local variations in the gray-level co-occurrence matrix.

$$contrast = \sum_{\mathrm{i,j}=0}^{N-1} \mathrm{P}_{ij}\,(i-j)^2 \tag{3}$$

Correlation is Measures the joint probability occurrence of the specified pixel pairs.

$$correlation = \sum_{\mathrm{i,j}=0}^{N-1} \mathrm{P}_{ij}\left(\frac{(i-\mu)(j-\mu)}{\sigma^2}\right) \tag{4}$$

Energy is Provides the sum of squared elements in the GLCM. Also known as uniformity or the angular second moment.

$$Energy = \sum_{\mathrm{i,j}=0}^{N-1} (\mathrm{P}_{ij})^2 \tag{5}$$

Homogeneity is Measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal.

$$Homogeneity = \sum_{\mathrm{i,j}=0}^{N-1} \left(\frac{\mathrm{P}_{ij}}{1+(i-j)^2}\right) \tag{6}$$

Where:
p_{ij} = Element i,j of the normalized symmetrical GLCM
N = Number of gray levels in the image as specified by Number of levels in under Quantization
μ = the GLCM mean (being an estimate of the intensity of all pixels in the relationships that contributed to the GLCM), calculate

$$\mu = \sum_{\mathrm{i,j}=0}^{N-1} i\mathrm{P}_{ij} \tag{7}$$

$\sigma =$ the variance of the intensities of all reference pixels in the relationships that contributed to the GLCM, calculated as:

$$\sigma^2 = \sum_{\mathrm{i,j}=0}^{N-1} \mathrm{P}_{ij}(i-\mu)^2 \tag{8}$$

From GLCM matrix we extract four features, and we added another two features to it. The two features are entropy and variance.
Entropy is a statistical measure of randomness that can be used to characterize the texture of the input image.

$$Entropy = \sum_{\mathrm{i,j}=0}^{N-1} -\ln(\mathrm{P}_{ij})\mathrm{P}_{ij} \tag{9}$$

Variance is a measure of variability. It is calculated by taking the average of squared deviations from the mean.

$$\sigma^2 = \frac{\sum(x-\mu)^2}{N} \tag{10}$$

Where:
χ is each value
$\mu =$ is the image mean
N is number of values in the image

The second group of features is the moment features [6,11]. It is a certain particular weighted average of the image pixels' intensities, or a function of such moments, usually chosen to have some attractive property or interpretation Based on normalized central moments, the invariant features can be achieved using central moments, which are defined as follows:

$$\mu_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\bar{x})^p(y-\bar{y})^q f(x,y)dxdy\text{p,q=0.1.2.}\ldots \tag{11}$$

Where:
$\bar{x} = (\frac{m_{10}}{m_{00}}), \bar{y} = (\frac{m_{01}}{m_{00}})$

The pixel point $(\bar{x}, \bar{y})$ are the centroid of the image f(x.y) of the image f(x.y) is equivalent to theμ_{pq},
whose center has been shifted to centroid of the image.
Therefore, the central moments are invariant to image translations.

$$\eta_{pq} = (\frac{\mu_{pq}}{\mu_{00}^{\gamma}}), \gamma = (\frac{p+q+2}{2})p+q = 2,3,.... \tag{12}$$

Scale invariance can be obtained by normalization. The normalized central moments are defined as follows:

$$\phi_1 = \eta_{20} + \eta_{02} \tag{13}$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \tag{14}$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \mu_{03})^2 \tag{15}$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \mu_{03})^2 \tag{16}$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}))[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \tag{17}$$

$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \tag{18}$$

$$\phi_7 = (3\eta_{21} + \eta_{03})(\eta_{30} + \eta_{12}))[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \tag{19}$$

From the moment features we extract seven features for each images. So the total number which extracted from each image is 13. This input features helped the classification model to classifying the images. But after extracting this features we applied first the discrete cosine transformation [7]. it is used in loss image compression because it has very strong energy compaction, i.e., its large amount of information is stored in very low frequency component of a signal and rest other frequency having very small data which can be stored by using very less number of bits (usually, at most 2 or 3 bit). To perform DCT Transformation on an image, first we have to fetch image file information (pixel value in term of integer having range 0 – 255) which we divide in block of 8×8 matrix and then we apply discrete cosine transform on that block of data.

Algorithm

1. Load the data set and store it.
2. Resize each image to make them in the same size (128×128).
3. Select the blue channel for the images, the data stores in 3D matrix.
4. Apply cosine transformation, and getting the coefficients as features.
5. Extract the moments, contrast, correlation, energy, and homogeneity.
6. Divide the features set into training and testing (80% for train and 20% for test)
7. Creating the model using the machine learning classifier with the training set.
8. Applying the model for classifying the test set.
9. Calculate the accuracy as equation (2)

The output accuracy value from this model in the first stage of training data is 74.63%, and the accuracy of the second stage of the model is 64.41%.

6.3. Third approach-linear regression classifier:

The Histograms of Oriented Gradients (HOG) features is another approach for feature extraction from images [4]. It is a widely used feature descriptor for object detection in computer vision tasks. Unlike edge features that only identify whether a pixel is an edge or not, the HOG descriptor considers the structure or shape of an object and provides the direction of the edge by calculating the gradient and orientation.The image is divided into smaller regions, and for each region, the gradients and orientation are calculated. This is known as ”localized” calculation. The HOG then generates a histogram for each of these regions, based on the gradients and orientation of the pixel values, hence the name ”Histogram of Oriented Gradients.”

In summary, the HOG feature descriptor counts the gradient orientation in localized portions of an image. The classification model used in this study is the multiple linear regression classifier [1].

Algorithm

1.Load the data set and store it
2.Image preprocessing which includes

a)Resize the images (128×128).
b)Extracting the hog features.
c)Store it in 3D Array ([128 128],1,N), and N is the number of samples.
3.Divide the data set into train and test data (80% for train and 20% for test)
4.Apply the linear regression classifier.
5.Calculate the accuracy as in equation (2)
The output accuracy value from this model in the first stage of training data is 98%, and the accuracy of the second stage of the model is 90%.

7. Analysis of the Results

The accuracy results from the three models indicate that the third model is the most effective. This model exhibits a high level of accuracy in classifying both binary classes (benign and malignant) and three classes (representing different types of cancer: carcinoma, sarcoma, and lymphoma). It is essential to emphasize that the selection of features significantly impacts the performance of the classifier. Multiple attempts were made to identify the most influential features for each classifier through a trial-and-error process, considering various combinations of features such as texture, shape, and color metrics.

Ultimately, the best combination of features and classifier was chosen, leading to the reported results. However, it is important to recognize that relying solely on visual information may not always yield the most precise outcomes. The integration of additional data sources, such as genetic markers or patient history, could enhance the accuracy further. Figure 5 and Table 1 present the comparison of accuracy ratios among the models.

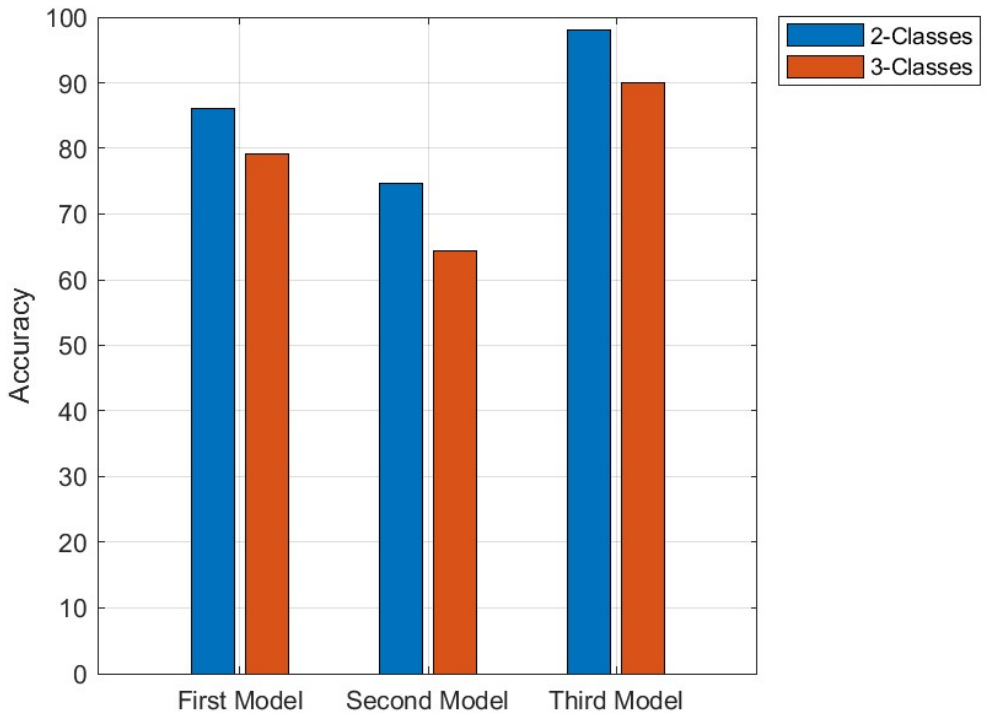

Figure 5. Fig (5): The accuracy results from the three models for the two stages.)

The analysis reveals the following:

• For the first model, the overall accuracy is calculated as 86.08% (accuracy for benign vs. malignant) multiplied by 79.12% (accuracy for the three cancer types), resulting in an overall accuracy of 68.11%.

• For the second model, the overall accuracy is 48.07%.

• For the third model, the overall accuracy is 88.20%.

The third model delivers the best performance in terms of accuracy.

Table 1. Comparison between The Three Models with respect to the Accuracy

Stages of the model	First approach	Second approach	Third approach
Stage 1	86.08%	74.63%	98%
Stage 2	79.12%	64.41%	90%
Overall	68.11%	48.07%	88.20%

Table 1 summarizes this analysis, showing that the third model approach achieves the highest accuracy in both stages: first, in the binary classification of benign versus malignant, and second, in the classification of detected cancer into one of the three primary cancer types. This stepwise classification approach ensures more precise identification and categorization of cancer types.

The third model stands out as the most reliable and effective, providing the highest accuracy for both stages of classification.

8. Conclusions

Skin cancer is on the rise and affects numerous individuals daily. Successful treatment of this cancer hinges on early detection. Early diagnosis and treatment not only increase the chances of survival but also reduce mortality rates. However, the current clinical techniques employed for diagnosing malignant melanoma are susceptible to human error due to subjectivity, the presence of novice physicians, and the high diagnostic costs. Consequently, there is a pressing need for more dependable and precise systems that can benefit both seasoned and inexperienced physicians, while remaining cost-effective for patients.

This paper introduces innovative and effective methods to aid in the visual interpretation of dermoscopy images, thereby reducing the occurrence of misdiagnoses. It proposes a novel approach for diagnosing and classifying skin cancer and its various types. The paper underscores that the accuracy results from the three models indicate the superiority of the third approach. This approach exhibits remarkable accuracy in classifying both two classes (benign and malignant) and three classes (representing the three types of cancer). It is crucial to recognize that the selection of features can significantly influence the classifier's performance. Numerous efforts were made to identify the most crucial features for each classifier, ultimately leading to the selection of the optimal features and classifier to yield these results.

REFERENCES

1. Adnan, M. N. B., Ahmad, W. M. A. W., Rahman, N. A., Ghazali, F. M. M., Alengc, N. A., Badrin, Z. M. Y., ... & Noor, N. F. M. *Arobust hybrid methodology between applied linear regression model (alrm) and multilayer perceptron (mlp).*, Bangladesh Journal of Medical Science, 22(1), 38-46.
2. Chen, X., Zhong, H., & Bao, Z., *A GLCM-Feature-Based Approach for Reversible Image Transformation.*, Computers, Materials & Continua ,(2019), 59(1).
3. Conte, S., Ghazawi, F. M., Le, M., Nedjar, H., Alakel, A., Lagacé, F., ... & Litvinov, I. V. *Population-Based Study detailing cutaneous melanoma incidence and mortality trends in Canada.*, Frontiers in medicine , 9, 830254,(2022).
4. Das, P., & Neelima, A, *A robust feature descriptor for biomedical image retrieval.*, IRBM, 42(4), 245-257. 2021.
5. Elgamal, M. , *Automatic skin cancer images classification.*, International Journal of Advanced Computer Science and Applications, 2013.
6. Huang, Z., & Leng, J., *Analysis of Hu's moment invariants on image scaling and rotation.*, IEEE international conference on computer engineering and technology (Vol. 7, pp. V7-476) (2010, April) .

7. Ince, I. F., Bulut, F., Kilic, I., Yildirim, M. E., & Ince O. F., *Low dynamic range discrete cosine transform (LDR-DCT) for high-performance JPEG image compression.*, The Visual Computer, 38(5), 1845-1870,(2022).
8. Janiesch, C., Zschech, P., & Heinrich, K., *Machine learning and deep learning.*, Electronic Markets, 31(3), 685-695, 2021.
9. Li, L., *Application of deep learning in image recognition.*, In Journal of Physics: Conference Series (Vol.1693,No.1, p.012128), 2020.
10. Naresh K. S., & Mohammed, I. B., *A novel approach to classify skin malignancy through deep convolutional neural network and image preprocessing approaches.*, International Journal of Health Sciences, 6(S1), 14194–14204, 2022.
11. Nasrudin, M. W., Yaakob, N. S., Rahim, N. A. A., Ahmad, M. Z. Z., Ramli, N., & Rashid, M. S. A., *Moment Invariants Technique for Image Analysis and Its Applications: A Review.*, In Journal of Physics: Conference Series (Vol. 1962, No. 1, p. 012028). IOP Publishing,(2021, July).
12. National Toxicology Program., *Ultraviolet-radiation-related exposures.*, In 15th Report on Carcinogens [Internet]. National Toxicology Program, 2021.
13. Nyemeesha, V. , *A systematic study and approach on detection of classification of skin cancer using back propagated artificial neural networks.*, Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(11), 1737-1748, 2021.
14. Nyemeesha, V., & Ismail, B. M., *Implementation of noise and hair removals from dermoscopy images using hybrid Gaussian filter.*, Network Modeling Analysis in Health Informatics and Bioinformatics, 2021.
15. Zabor, E. C., Reddy, C. A., Tendulkar, R. D., & Patil, S., *Logistic regression in clinical studies.*, International Journal of Radiation Oncology, Biology, Physics, 112(2), 271-277, 2022.