Improved Model for Skin Illnesses Classification Utilizing Gray-Level Co-occurrence Matrix and Convolution Neural Network

Zahraa Jabbar Hussein [1,*], Ahmed Mohammed Hussein [1], Ghadeer Ibrahim Maki [2], and Hadeel Qasem Gheni [1]

[1] Department of Computer Science, College of Science for Women, University of Babylon, Babylon, Iraq; Email: wsci.ahmed.mohammed@uobabylon.edu.iq (A.M.H.), wsci.hadeel.qasem@uobabylon.edu.iq (H.Q.G.)
[2] Nursing Techniques Department, Technical Institute of Al-Diwaniyah, Al-Furat Al-Awsat Technical University, Najaf, Iraq; Email: Ghdeer.maki.idi@atu.edu.iq (G.I.M.)
*Correspondence: zahraa.jabbar@uobabylon.edu.iq (Z.J.H.)

***Abstract*—Skin conditions are more common than other illnesses. Skin issues can be brought on by viruses, bacteria, allergies, fungi, etc. The detection of skin diseases has been made better by lasers and photonics since it is now quicker and more precise. However, such a diagnosis is pricey. A system for automated dermatology screening is built with the use of computer vision. Using the Gray Level Co-occurrence Matrix (GLCM) and Convolution Neural Network (CNN) provides an improved model for accurately diagnosing skin disorders. In order to classify skin photos, CNN is used by the model to extract features from the images using GLCM. The high-level features utilizing the statistical features were retrieved via GLCM separately because the photos utilized in the research are for skin conditions. Once merged, these features created a high-accuracy categorization. Two distinct classification processes are used to categorize photos into 13 diseases: First, the Deep Neural Network (DNN) classifier obtains 96.69% accuracy, 96.2% recall, 96.2% precision, and 96.2% F1-score in terms of performance evaluation measures. Second, accuracy, recall, precision, and F1-score are the performance evaluation metrics for the Multiple Support Vector Machine (MSVM) classifiers. The model outperforms other cutting-edge models in terms of accuracy and effectiveness when compared to them. This work thus indicates the capability of GLCM and CNN for the classification of skin diseases and their prospective uses in the healthcare sector.**

***Keywords*—deep learning, convolution neural network, gray-level co-occurrence matrix, skin disease**

I. Introduction

Skin conditions are a major global healthcare concern that impact people of all ages and socioeconomic groups. Numerous diseases that affect the skin, hair, nails, and mucous membranes are included in these ailments. The burden that skin illnesses place on healthcare systems and the prevalence, complexity, and impact of these conditions on people's quality of life are the issues [1]. Skin conditions can present in a variety of ways, from minor irritations and rashes to chronic illnesses and possibly fatal conditions. They can have an impact on people regardless of their age, gender, or race. These disorders may be brought on by infectious agents such as bacteria, viruses, fungi, or parasites, autoimmune reactions, environmental triggers, or hereditary factors [1, 2]. Around the world, between 30% and 70% of people suffer from skin illnesses, which can afflict anyone at any age or at any point on the body. Our investigation was prompted by numerous issues, including [3, 4]:

- Only 24% to 70% of diagnoses made by specialists are accurate. Due to delayed or ineffective treatment, their lack of experience and training in a field with a large number of cases may cause the sickness to worsen and progress.
- Skin illnesses can exhibit a variety of symptoms and can develop similarly, making proper diagnosis challenging. Differentiating between diverse skin problems requires specific knowledge and skills, which could result in misdiagnosis or delays, compromising the ability to receive prompt and effective treatment.
- In some areas or for marginalized groups, access to dermatological care and treatment may be restricted, leading to inadequate or delayed maintenance. Additionally, dermatological operations, drugs, and treatments can be costly, placing a strain on people's finances, especially those who lack proper health insurance.

There has been no action done to address these skin disorders despite their high prevalence and the resulting poor diagnostic precision of experts. Dermatological diseases can now more easily be identified using merely an image thanks to developments in the science of artificial intelligence and associated technologies [5]. Skin illness detection is fraught with difficulties [3–6]:

- Due to hair, perspiration, and other properties, human skin is one of the most intricate and unexpected regions that may be automatically evaluated and synthesized.

Manuscript received March 12, 2023; revised June 25, 2023; accepted July 19, 2023; published November 27, 2023.

- Due to the visual similarity of many skin problems and the intricacy of skin texture, it is challenging for computer vision to distinguish between different dermatological from colour photos. But this work presents a huge obstacle.
- Digital skin images have many issues, including noise, quality, and inconsistent lighting.
- Skin disease diagnosis is made more challenging by lesion impurities such skin tone, hair colour, air bubbles, and nails.
- It could be challenging to recognize the specific type of skin issue due to the visible closeness of illnesses and the intricacy of skin texture.

The proposed approach was used in the current paper to handle 13 different forms of skin disorders. using two feature extraction methods: machine learning, which provides statistical features separate from Convolution Neural Network (CNN) features, and deep learning, which is now the most popular. The categorization of the more than 10 diseases utilized in this work differs from the research that will be covered in the second half, as does the model of the proposed combination and classification using the first two techniques, Deep Neural Network (DNN) and Multiple Support Vector Machine (MSVM). As can be seen in the Results and Discussion section, the system outperformed pre-made CNN models.

An overview of the paper presented in this publication is given below. The description of related works in Section II follows the introduction to this study in Section I. Data collecting is covered in Section III, along with data augmentation, CNN, Gray Level Co-occurrence Matrix (GLCM), and classifiers using DNN and Support Vector Machine (SVM). In Section IV, the experimental inquiry is described. The conclusions are reported in Section V.

II. Literature Review

Recent studies conducted in this area have demonstrated that the use of artificial neural networks in the medical field aids decision-making and that the automated diagnosis of skin illnesses is based on a decision-making system for skin diseases. Here, we look at a few research projects in this area. Based on the style-based GANs architecture, a skin lesion style-based Generative Adversarial Networks (GANs) model is developed [6]. It produces high-resolution, varied images of skin lesions that are effective. Transfer learning is used to build the deep transfer neural network-based classifier. To enhance the training set, the recommended GAN-based models' synthetic images were added for ISIC 2018 dataset [7] in order to improve classification performance.

A model with four phases—image processing, segmentation of skin lesions, feature extraction, and DNN-based classification [8]. Image processing eliminates noise, Otsu's image segmentation is used for segmentation, and GLCM, 2D Discrete Wavelet Transform (DWT), and RGB color models are used to extract features. The proposed deep learning model was trained and tested using the ISIC 2017 dataset. It achieved 84.45% accuracy with DNN [8].

Segmentation, feature extraction, and latent representation were used in a novel detection framework to enhance the diagnosis of skin diseases and outperform traditional deep classification techniques [9]. CNN, GWO, GLCM, WLD, and AutoEncoder are employed to categorize pathogenic lesions from latent representation. The obtained accuracy is 94.2% [9].

Pandu in 2023 utilized GLCM, 2D DWT, and RGB color models are to image segmentation, extract features, and image processing removes noise. The ISIC 2017 dataset was used to train and test the suggested deep learning model. With DNN, it obtained 84.45% accuracy. K-means clustering, an unsupervised learning technique, was used by Akar to successfully pinpoint the area of the skin that is impacted by cancer. Then, using the Gray-Level Co-occurrence Matrix (GLCM), the Discrete Wavelet Transform (DWT), and computation of image statistics for texture, low-level features, and color features, respectively, many characteristics are recovered from the segmented output. In order to classify skin cancer with the highest degree of efficiency, the Probabilistic Neural Network (PNN) was finally developed using, respectively, statistical color characteristics, low-level features based on the Discrete Wavelet Transform (DWT), and texture characteristics based on the Grey-Level Co-occurrence Matrix (GLCM) [10].

Using CNN architectures, MobileNet, and Xception, a computer vision-based application that can recognize different skin conditions was created [10]. The suggested CNN architectures employed transfer learning to find additional features by pre-training models on the ImageNet dataset. Additionally, we assessed how well our proposed strategy performed using some of the most popular CNN designs, including ResNet50, InceptionV3, Inception-ResNet, and DenseNet, to create a standard that would support the fundamental principles of augmentation and transfer learning. 97.00% accuracy was attained [11].

For the purpose of categorizing and detecting images, a model utilizing SVM, K-Nearest Neighbors (KNN), and random forest algorithms was developed [12]. Image analysis to reliably identify skin diseases is made easier with the presented methodology. The accuracy percentage for the SVM approach was 98.8%. The specificity and sensitivity meatures were 99% and 91%, respectively, when classifying using KNN algorithm [12].

III. Research Methods

This section discusses the following procedures: describe the data set and methodology, including (1) data augmentation, (2) feature extraction with CNN and GLCM, and (3) classification using DNN and MSVM.

This article suggested enhancing diagnostic precision and expanding accessibility to dermatology knowledge. The paper examines how recent developments in deep learning have made it possible to develop tools that can help with image-based skin condition diagnosis. The objective is to make skin disease diagnostic support more readily available and accurate, especially for individuals who might not have access to dermatology clinics or dermatologists who specialize in skin disorders. The classification of people's skin conditions is the goal of the suggested system. Finding the traits of each kind to

distinguish between them is necessary because many diseases share a similar outward appearance, making it difficult for such computerized systems to recognize and classify the type of sickness. Two methods are used to extract the characteristics: CNN deep learning, which extracts high-dimensional features, and machine learning, which uses one of the tissue algorithms to identify between the tissues of the diseases depicted in the image. These features are merged into a single vector after being derived using CNN and GLCM. Merge approach via a series of actions:

- Standard division is used to normalize the data, bringing it inside a predetermined range and unifying its value.
- Utilizing Singular Value Decomposition (SVD), this vector's features are reduced, resulting in fewer features. Finally, classification algorithms are used to classify these features.

A. Data Set

There are a total of 13 skin diseases listed in this database, and they are as follows [13]: Atopic Dermatitis, Acne, Rosacea, Bullous Disease, Cellulitis Impetigo, Eczema, Herpes HPV, Vascular Tumours, Psoriasis Onycholysis, Pigmentation Disorders, Melanoma Skin Cancer Ringworm Candidiasis in Nevi and Tinea. Samples from the dataset are shown in the first figure. There are 11,894 data in total. The number of photos varies depending on the type of skin illness. The data was split into two sections: training, with a total of 9,416 records, and testing, with a total of 2,478 records. The distribution of the dataset is displayed in the Table I below.

TABLE I. EXPLAIN THE NUMBER OF IMAGES IN THE DATASET

Type of Disease	Total of images	No. of train	No. of Test
Atopic Dermatitis	612	489	123
Acne and Rosacea	1,152	840	312
Bullous Disease	561	448	113
Actinic Keratosis Basal Cell Carcinoma	1,437	1,149	288
Cellulitis Impetigo	301	228	73
Exanthems	505	404	101
Eczema	1,544	1,235	309
Herpes HPV	507	405	102
Vascular Tumors	603	482	121
Psoriasis Onycholysis	1,757	1,405	352
Disorders of Pigmentation	711	568	143
Melanoma Skin Cancer Nevi	579	463	116
Tinea Ringworm Candidiasis	1,625	1,300	325

B. Data Augmentation

In order to train discriminative deep learning models, additional data must be provided. A large amount of labeled training data is necessary for convolutional neural networks. Because CNN's tuning is subpar, a small training set will lead to overfitting. It costs money to get enough medical images. Consequently, this tactic will be applied. By rotating, turning, resizing, and sharing photographs, data augmentation was used to enhance results. Rotation has no effect on the model's orientation. Rotate images at random between 0 and 360 degrees. When an image is rotated, some of the pixels move, necessitating a fill-in with the reflected image. The image is horizontally flipped and then randomly zoomed in. Shearing is a technique used to enhance photographs so that computers can recognize various viewpoints. They randomly move an image up, down, left, or right [14, 15] (see Fig. 1).

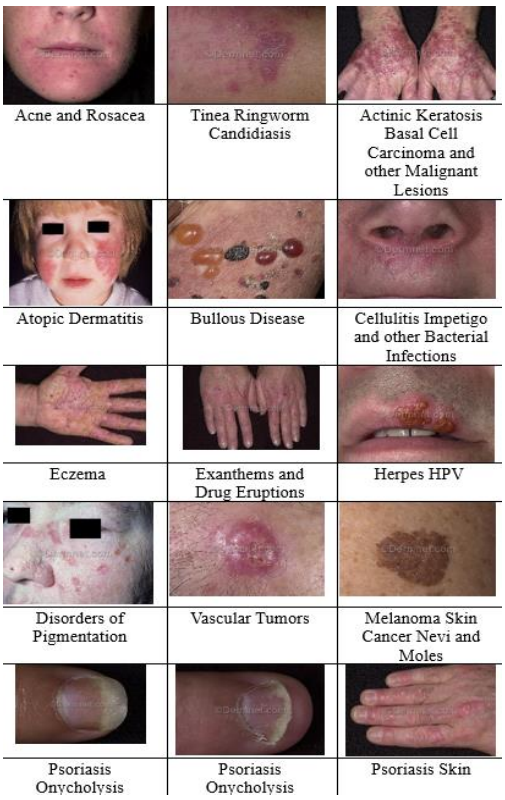

Figure 1. Sample of dataset.

C. Convolution Neural Network

DNN-based machine learning, or deep learning, has hidden DNN layers. Feature extraction is automatic, unlike machine learning. A gradient-based DNN is CNN. When feature-extractor and classifier networks are combined, a hidden-layer DNN can show how the brain classifies images [16]. There are layers of convolution and pooling in the feature-extractor network. Multi-class classifier neural networks are used. Extraction of CNN features prior to training. CNN's pre-training model design is shown in Table II.

- Rows, columns, and layers are all defined by the input layer for the input dimensions (image).
- Convolution layers: To produce the features map, this layer generates input data and applies filters to it. Filter number, size, and user padding are among the parameters of this layer; the filtering algorithms' weights are updated always in every period.
- ReLU AC: Since the negative number does not reflect features after convolution, it is ignored. Eq. (1) explains the ReLU function.

$$(\mathcal{S}) = \max(0, \mathcal{S}) = \begin{cases} \mathcal{S}_i, if\ \mathcal{S}_i \geq 0 \\ 0, if\ \mathcal{S}_i < 0 \end{cases} \quad (1)$$

In Eq. (1), (*S*) stands for a variable or a group of values. The greatest value between 0 and *S* is produced by the function *f*(*S*). The output is similar to S itself if the value of *S* is larger than or equal to 0. However, if *S* has a negative value, the outcome is 0.

In particular, machine learning and neural networks frequently use the ReLU function. It adds nonlinearity to the model by acting as an activation function. The ReLU function aids in selectively stimulating some neurons while inhibiting others by putting negative values to zero.

The ReLU function's ease of use and effectiveness as a computational tool are two of its key benefits. Compared to other activation functions like sigmoid or tanh, it enables quicker training of neural networks. The vanishing gradient issue, which can arise during backpropagation in deep neural networks, is another issue that the ReLU function helps to solve.

In conclusion, computing efficiency was increased by using Eq. (1).

- Batch-normalization: By normalizing the input data, it is hoped to increase the reliability and effectiveness of neural network training. They are often used following the end of each convolutional layer.
- Max-Pooling: The network produces feature maps following convolution. The lowest and best versions of these qualities are left in this layer. Feature maps are made up of 2×2 blocks. The reduction employs the block average or the most significant block value along with half-sized features.
- Dropout: Dropout lowers overfitting and raises the generalizability of the model. Depending on their likelihood of being in the chosen hidden layers, this technique ignores hidden neurons [17]. A neural network's dropout is depicted in Fig. 2.

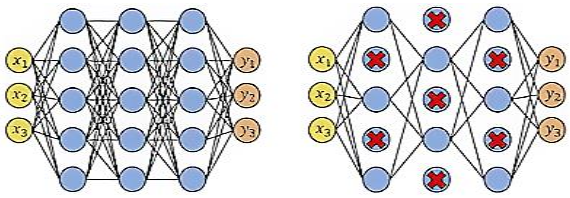

Figure 2. Dropout in neural network.

- There are numerous levels (convolution, activation function, pooling, and finally flatten a layer.
- Soft-max: In the classification stage, this function serves as an activation function for classifying several classes. Eq. (2) explains the SoftMax function.

$$f(x) = \frac{1}{1+e^{-x}} \tag{2}$$

A vector of real values is transformed into a probability distribution using Eq. (2). In issues involving several classes of categorization, it distributes probabilities to various classes. The SoftMax function exponentiates and normalizes the input numbers to guarantee that the possibilities that result add to 1. For multi-class classification problems, it is frequently employed as the last activation function in neural networks, enabling the model to make predictions based on class probabilities. Accurate forecasts in multi-class scenarios are made possible by the SoftMax function, which assigns higher probabilities to more pertinent classes and suppresses the likelihood of less likely classes (see Table II).

TABLE II. THE SYSTEM'S SUGGESTED CNN MODEL'S ARCHITECTURE

Name	Type	Activation
Image input	Image input	250 × 250 × 3
Conv_1	2-D Conv.	250 × 250 × 30
Relu_1	Activation Fun.	250 × 250 × 30
BatchNo_1	Batch norm.	250 × 250 × 30
Maxpool_1	2-D max Po.	125 × 125 × 30
Conv_2	2-D Conv.	125 × 125 × 60
Relu_2	Activation Fun.	125 × 125 × 60
BatchNo_2	Batch norm.	125 × 125 × 60
Maxpool_2	2-D max Po.	62 × 62 × 60
Conv_3	2-D Conv.	62 × 62 × 90
Relu_3	Activation Fun.	62 × 62 × 90
BatchNo_3	Batch norm.	62 × 62 × 90
Maxpool_3	2-D max Po.	31 × 31 × 90
Conv_4	2-D Conv.	31 × 31 × 30
Relu_4	Activation Fun.	31 × 31 × 30
BatchNo_4	Batch norm.	31 × 31 × 30
Maxpool_4	2-D max Po.	15 × 15 × 30
Conv_5	2-D Conv.	15 × 15 × 60
Relu_5	Activation Fun.	15 × 15 × 60
BatchNo_5	Batch norm.	15 × 15 × 60
Maxpool_5	2-D max Po.	7 × 7 × 30
FC	Fully Conn.	1 × 1 × 1500
Drop.	-	0.5
Sigmoid_1	Activation Fun.	-
Class output	Classification	13
Option	Training Opt.	Optimizer = Adam Epoch=90, learn rate=0.00001

D. Grey-Level Co-occurrence Matrix Technique

Using the grey-level co-occurrence matrix, the discovered skin's textural characteristics have been retrieved. The method of reducing data to find a select group of practical image-based variables called feature extraction. Six textural components are extracted from each image in this study using the Grey-Level Co-occurrence Matrix (GLCM). Co-occurrence matrices are created for angles of 0, 45, 90, and 135 degrees. From the co-occurrence matrices produced for each of the four angles, the seven Haralick texture descriptors are obtained [18, 19].

- Entropy: When paired with texture analysis, it measures the disturbance in the spatial space and is an indication of discontent in any given system. It may have been calculated as in Eq. (3):

$$En_t = -\sum_{p=0}^{N-1}\sum_{k=1}^{N-1} Gl(p,k)\, logGl(p,k) \tag{3}$$

Accurate forecasts in multi-class scenarios are made possible by the SoftMax function, which assigns higher probabilities to more pertinent classes and suppresses the likelihood of less likely classes. By reducing entropy or cross-entropy loss during training, models develop the ability to make predictions with more assurance and

accuracy. Entropy is a measurement parameter used in generative models to evaluate the variety or uncertainty of generated samples. It is essential for deep learning network training and for assessing how well models perform in classification or generative modeling tasks.

- Contrast: relates to the variety of grayscale values. You can figure it out by:

$$Co_n = \sum_{p=0}^{N-1} \sum_{k=1}^{N-1} (p-k)^2 \, Gl(p,k) \tag{4}$$

Eq. (4) uses the term "contrast enhancement" to describe methods that increase the visibility and distinctions between features in a dataset. A common method for redistributing intensity values to increase contrast is histogram equalization. Contrast enhancement methods preprocess data and boost the discriminative ability of features for deep learning models by amplifying the contrasts between significant patterns and background noise. These methods are helpful for boosting feature extraction, extracting subtle patterns, and improving texture details in a variety of data sources.

- Energy: Entropy measures global heterogeneity; energy measures local homogeneity. This characteristic shows how uniform the texture is. You can figure it out by:

$$En_r = \sum_{p=0}^{N-1} \sum_{k=1}^{N-1} Gl(p,k)^2 \tag{5}$$

Eq. (5) in texture analysis adds up the squared values of the components of a Grey-Level Co-occurrence Matrix (GLCM) to determine the energy of a texture. gives a measurement of the texture's local homogeneity or uniformity. Greater homogeneity is indicated by higher energy levels, whilst more variable or heterogeneous textures are indicated by lower energy values. In contrast to entropy, the energy component emphasizes the uniformity rather than the intricacy of the texture. It is frequently used to evaluate and identify the uniformity of textures in activities including texture analysis, pattern identification, and image processing.

- Dissimilarity: It describes the various grey-level pairs that make up an original image. It might be calculated as:

$$Di_s = \sum_{p=0}^{N-1} \sum_{k=1}^{N-1} |p-k| \, Gl(p,k) \tag{6}$$

Eq. (6) determines the degree of variance or distinctions between pairs of grey levels in an original image. The dissimilarity function, which distributes weights to the squared disparities between grey-level values over all pixel pairings in the image, is added up in this process. The statistical link between the pairs of grey levels is captured by the joint intensity function. The resulting measure of dissimilarity tells us how varied or distinctive the grey-level associations in the image are. The diversity and uniqueness of grey-level pairs are frequently evaluated in texture classification, pattern recognition, and picture similarity comparison tasks.

- Homogeneity: It speaks to the GLCM's nonzero entries' consistency. It might be calculated as:

$$Di_s = \sum_{p=0}^{N-1} \sum_{k=1}^{N-1} \frac{1}{1-(p-k)^2} \, Gl(p,k) \tag{7}$$

Eq. (7) in a Gray-Level Co-occurrence Matrix (GLCM), homogeneity quantifies how consistently or uniformly nonzero items appear. measures how much an image's spatial relationships between adjacent pixels display predictable patterns. The squared absolute difference between the intensity values of the pixel pairings is divided by the total of the GLCM entries in the homogeneity equation, which then yields the homogeneity value. Greater uniformity in the visual texture is indicated by higher homogeneity levels. For the purpose of distinguishing between various texture patterns and spotting anomalies in medical imaging, homogeneity is frequently utilized in image analysis and texture classification tasks.

- Correlation: It shows how the texture of the image resembles it in two different directions, notably in the *x* and *y* directions. It might be calculated as:

$$cor_r = \frac{\sum_{p=0}^{N-1} \sum_{k=1}^{N-1} (p-m^p)(k-m^k) Gl(p,k)}{\sigma_p \sigma_k} \tag{8}$$

Eq. (8) calculates how well two different directions of an image's texture match. The correlation coefficient is calculated by adding the variations in pixel intensities in each direction from their corresponding means, weighted by the combined intensity of the pixel pairs. To normalize the covariance, the result is divided by the sum of the standard deviations in each direction. The linear link between the two directions of the texture is indicated by the correlation coefficient, along with its intensity and direction. It is utilized in jobs involving texture analysis and pattern identification to evaluate how similar or distinct textures are and to find enduring patterns in various orientations.

The data from the photographs is entered into the GLCM method once these data images have been converted from RGB to grey. The features that will be coupled with the features that came from the pre-training model are then extracted using a GLCM matrix. These properties include Entropy, Contrast, Energy, Dissimilarity, Homogeneity, and Correlation.

E. Classification

The first type of classification employed in this model is DNN, while the second type is MSVM [20]. A single collection of features that combines the output of the CNN model and the GLCM method is utilized as the input to the DNN or the SVM. The DNN is made up of two layers, including two input layers (GLCM and CNN features), a concatenate layer that combines the two levels below it, three dense layers, and the ReLU activation function that follows right away. After each dense layer, two dropouts with a different probability of 0.5 will be applied. These

dropouts are there to avoid overfitting. A SoftMax activation function-containing output layer is used to calculate the likelihood of class 13 disorders.

IV. Results and the Discussion

The results obtained from the suggested model are examined in this section. It's crucial to separate the data into training and testing sets before building a model. It is required to divide the data into training and testing sets in order to evaluate the model's performance on the brand-new, unexplored data. Depending on the size of the dataset, the ratio of the training set to the testing set may change. For the training sets, the standard ratio is 80%, and for the testing sets, 20%. To prevent any biases in the model, it is crucial to make sure that the data are separated at random. Using a random number generator to divide the data is one technique to guarantee that the division is random. It is necessary to apply various criteria to each suggested system in order to determine its accuracy, precision, recall, and F1-score [11, 12]. Test data used to validate the system is used to determine the values for these metrics.

One of the most crucial metrics in machine learning is accuracy because it gauges the model's quality. This action Lacks deal with data imbalance problems in datasets. From Eq. (9), the accuracy value may be derived.

$$Accuracy = \left[\frac{TP+TN}{TP+TN+FP+FN} \times 100\right] \quad (9)$$

where True Positive (TP): The real class is true and the model predicates true because it predicted positively and properly identified the classes.

- True Negative (TN): The model properly identified and predicted a negative outcome, which indicates that the model's prediction of the outcome was incorrect.
- False Positive (FP): The actual class is invalid, but the model is true since it is predicated improperly.
- False Negative (FN): The model is wrong and the actual class is true since the model's predicate is flawed.
- Eq. (10), which calculates the ratio of accurately positive predicted observations to all positive predicted observations, is used to determine precision.

$$Precision = \left[\frac{TP}{TP+FP} \times 100\right] \quad (10)$$

The proper ratio of correctly positive projected observations to all actual observations in the class is used to calculate recall or sensitivity. You can get it by using Eq. (11).

$$Recall = \left[\frac{TP}{TP+FN} \times 100\right] \quad (11)$$

The F1-Score is determined by using false positive and false negative results to weight the weighted mean of precision and recall. In general, accuracy performs best when false positives and negatives have comparable prices, whereas the F1-score performs best when the value of a false positive differs from a false negative. It is preferable in this situation to look at precisions and recall when the dataset has an unequal class distribution. The value of it can be determined using Eq. (12).

$$F_1 = \frac{2(Precision \times Recall)}{Precision + Recall} \times 100 \quad (12)$$

Test results are categorized in accordance with the optimum state discovered during training. By dividing the total number of correctly diagnosed test data by the total quantity of test data, we may determine how accurate the system is. Number of correctly or mistakenly diagnosed data yields precision, recall, and F1-score.

DNN and MSVM, two of the most well-known and effective classification methods, were employed to categorize various data kinds. Our model's accuracy with DNN technology was 96.69%, the maximum accuracy compared to MSVM, with a very small difference of 96.44% according to Table III.

TABLE III. Performance Metrics Our Model

Model	Acc. (%)	Recall (%)	Precision (%)	F1-score (%)
DNN	96.69	96.2	96.2	96.2
MSVM	96.44	96.52	96.9	96.9

The other measurements, recall, precision, and F1-score, are comparable to the accuracy value, on the other hand. The recall ratio, precision, and F1-score values for the MSVM classifier were 96.52%, 96.9%, and 96.9%, respectively, while these values were 96.2%, 96.2%, and 96.2% for the DNN classifier. As a result, the value of these three measures for the MSVM classifier was greater than that of the DNN classifier.

The proposed model has a classification accuracy of 96.69% for skin conditions. As demonstrated in Table IV, this accuracy rate is much greater than that of other cutting-edge models like Alexnet, VGG16, VGG19, and GoogleNet.The accuracy in these models varied from 91% to 94%. The more accurate the diagnosis, which has an impact on the patients' life and is a medical study, the more accurate the study must be. Among these models, GoogleNet had the highest accuracy (94.4%), followed by Alexnet (93.5%), VGG19 (92.14%), and VGG16 (91.11%). Figs. 3–6 show how the suggested model and other models performed differently in terms of the accuracy of the scales.

TABLE IV. Compression between our Models and Others

Model	Acc. %	Recall %	Precision %	F1-score %
Alexnet	93.5	93.15	93.20	93.11
VGG16	91.11	91	91	91
VGG19	92.14	92.4	92.2	92.2
GoogleNet	94.9	94.4	94.3	94.4
Proposed +DNN	96.69	96.2	96.2	96.2
Proposed +MSVM	96.44	96.52	96.9	96.9

Our research demonstrates that the performance of classifying skin diseases can be enhanced by using GLCM as a feature extraction technique and CNN as a classification model. Additionally, we ran tests to assess how the suggested model performed in various settings. The findings show that, even with noisy or insufficient

data, the suggested model is reliable and efficient at classifying skin disorders. The great accuracy of the findings of the proposed model, when applied to the dataset presented in Section III implies that the extracted features depend on the statistical properties that depend on the image texture. Since deep learning, particularly when combined with CNN, has become quite important in categorization. Because dermatology images depend on the texture of the epidermis, utilizing such methods to extract characteristics has a significant impact on improving diagnosis accuracy. The potential of GLCM and CNN for classifying skin diseases is demonstrated in the current paper, which has important ramifications for the healthcare sector. The proposed methodology can help physicians diagnose skin disorders accurately and quickly, leading to better patient outcomes.

V. Conclusion

Skin disease detection can slow down disease progression, disease transmission, and death rates. Clinical tests that are expensive and time-consuming find skin conditions. A dermatological screening system is built with the aid of early picture processing. Skin conditions are classified using features. This research led to the creation of an enhanced method that combines CNN and GLCM with two classifications (DNN, MSVM). The experimental results of this model's application to extracting characteristics and classifying 13 different diseases revealed a significant improvement in classification accuracy when compared to earlier studies that divided diseases into 2–10 kinds. The model's evaluation metrics made it very clear that the accuracy attained was 96.69% for the DNN classifier and 96.44% for the MSVM classifier. The strength of the suggested work in categorizing diseases is evident from the results of additional metrics including recall, precision, and F1-score. Due to the significant degree of resemblance among diseases, an estimated 3.9% of misdiagnoses were discovered. Other skin problems can be categorized using this in more detail. The efficacy and patient care of dermatologists are improved by this work. To improve the findings, this model can be enhanced by adding more data.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

The database was prepared and the pre-processing was done by Hadeel. The paper was prepared in terms of writing and typesetting by Hadeel and Ghadeer. The software part was written by Zahraa and Ahmed. All authors had approved the final version.

References

[1] W. C. C. Y. Lu, M. S. Hsieh, K. C. Wei, M. Ezmerli, and H. Kuo, "Gastrointestinal involvement of primary skin diseases," *J. Eur. Acad. Dermatology Venereol.*, vol. 36, no. 12, pp. 1468–3083, 2022.

[2] E. I. Z. D. A. Gavrilov, A. V. Melerzanov, and N. N. Shchelkunov, "Use of neural network based deep learning techniques for the diagnostics of skin diseases," *Biomed. Eng.*, vol. 52, no. 5, pp. 348–352, 2019.

[3] M. N. Bajwa *et al*., "Applied sciences computer-aided diagnosis of skin diseases using deep neural networks," *Appl. Sci. Artic*., vol. 10, no. 2488, pp. 1–13, 2020.

[4] D. Filimon and A. Albu, "Skin diseases diagnosis using artificial neural networks," *IEEE Int. Symp. Appl. Comput. Intell. Informatics*, pp. 189–194, 2014.

[5] J. E. C. Drerup, "Autoinflammatorische," *Der Hautarzt*, vol. 71, pp. 359–364, 2020.

[6] H. Wu *et al*., "A deep learning, image-based approach for automated diagnosis for inflammatory skin diseases," *Ann. Transl. Med*., vol. 8, no. 9, pp. 1–8, 2020.

[7] H. Zhou, F. Xie, and Z. Jiang, "Multi-classification of skin diseases for dermoscopy images using deep learning," *IEEE Instrum*., pp. 18–22, 2017.

[8] N. Soliman and A. Alenezi, "A method of skin disease detection using image processing and machine learning," *Procedia Comput.* Sci., vol. 163, pp. 85–92, 2019.

[9] S. H. Bhadul *et al*., "Machine learning algorithms based skin disease detection," *International Journal of Innovative Technology and Exploring Engineering*, pp. 4044–4049, 2019.

[10] T. Shanthi, R. S. Sabeenian, and R. Anand, "Microprocessors and microsystems automatic diagnosis of skin diseases using convolution neural network," *Microprocess. Microsyst.*, vol. 76, 103074, 2020.

[11] O. P. Yanchatuña *et al*. "Skin lesion detection and classification using convolutional neural network for deep feature extraction and support vector machine," *International Journal on Advanced Science Engineering Information Technology*, vol. 11, 2021.

[12] A. V. Reddy, "A machine learning model for skin disease classification using convolution neural network," *International Journal of Computing, Programming, and Database Management*, vol. 3, no. 1, 2022.

[13] P. Kavitha, V. Jayalakshmi, and S. Kamalakkannan, "Classification of skin cancer segmentation using hybrid partial differential equation with fuzzy clustering based on machine learning techniques," in *Proc. 2022 International Conference on Edge Computing and Applications*, 2022.

[14] A. Esteva, K. Chou, S. Yeung, N. Naik, A. Madani, A. Mottaghi, Y. Liu, E. Topol, J. Dean, and R. Socher, "Deep learning-enabled medical computer vision," *NPJ Digit. Med.*, vol..4, pp. 1–9, 2021.

[15] L. Alzubaidi, M. A. Fadhel, S. R. Oleiwi, O. AlShamma, and J. Zhang, "DFU_qutnet: Diabetic foot ulcer classification using novel deep convolutional neural network," *Multimed. Tools Appl*., vol. 79, pp. 15655–15677, 2020.

[16] L. Alzubaidi, A. A. Abbood, M. A. Fadhel, O. AlShamma, and J. Zhang, "Comparison of hybrid convolutional neural networks models for diabetic foot ulcer classification," *J. Eng. Sci. Technol*., vol. 16, pp. 2001–2017, 2021.

[17] I. S. Razaq and B. K. Shukur, "Improved face morphing attack detection method using PCA and convolutional neural network," *Karbala International Journal of Modern Science*, vol. 9, no. 2, 2023.

[18] K. Pal and B. V. Patel, "Data classification with k-fold cross validation and holdout accuracy estimation methods with 5 different machine learning techniques," in *Proc. 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)*, 2020, pp. 83–87.

[19] A. C. I. Ossai and N. Wickramasinghe, "GLCM and statistical features extraction technique with extra-tree classifier in macular oedema risk diagnosis," *Biomed. Signal Process. Control*, vol. 73, 103471, 2022.

[20] K. Rangarajan and R. Purushothaman, "Disease classification in eggplant using pre-trained interests include artificial intelligence, machine learning, and data mining," *VGG16 and MSVM. Sci. Reports*, vol. 10, pp. 1–11, 2020.