IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 17 June 2023, accepted 5 July 2023, date of publication 13 July 2023, date of current version 19 July 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3294974

RESEARCH ARTICLE

Performance Enhancement of Skin Cancer Classification Using Computer Vision

AHMED MAGDY[1], (Member, IEEE), HADEER HUSSEIN[1], REHAB F. ABDEL-KADER[2], AND KHALED ABD EL SALAM[1]

[1]Electrical Engineering Department, Suez Canal University, Ismailia 41522, Egypt
[2]Electrical Engineering Department, Port Said University, Port Said 42523, Egypt

Corresponding author: Rehab F. Abdel-Kader (rehabfarouk@eng.psu.edu.eg)

ABSTRACT Nowadays, computer vision plays an essential role in disease detection, computer-aided diagnosis, and patient risk identification. This is especially true for skin cancer, which can be fatal if not diagnosed in its early stages. For this purpose, several computer-aided diagnostic and detection systems have been created in the past. They were limited in their performance because of the complicated visual characteristics of skin lesion images, which included inhomogeneous features and hazy borders. In this paper, we proposed two methods for detecting and classifying dermoscopic images into benign and malignant tumors. The first method is using k-nearest neighbor (KNN) as classifier when pretrained deep neural networks are used as feature extractors. The second one is AlexNet with grey wolf optimizer, that optimizes AlexNet's hyperparameters to get the best results. We also tested two approaches in classifying skin cancer images, which are machine learning (ML) and deep learning (DL). The used methods in ML approach are artificial neural network, KNN, support vector machine, Naïve Bayes, and decision tree. The DL approach that we used contains convolutional neural network and pretrained DL networks: AlexNet, VGG-16, VGG-19, EfficientNet-b0, ResNet-18, ResNet-50, ResNet-101, DenseNet-201, Inception-v3, and MobileNet-v2. Our experiments are trained and tested on 4000 images from the ISIC archive dataset. The outcomes showed that the proposed methods outperformed the other tested approaches. Accuracy of first proposed method exceeded 99% in some models and second proposed method achieved 99%.

INDEX TERMS Deep learning, machine learning, melanoma (malignant), nonmelanoma (benign), skin cancer.

I. INTRODUCTION

In the present decade, skin cancer stands out as one of the most prevalent forms of cancer [1]. This is not surprising considering that the skin serves as the largest organ in the human body, making it naturally susceptible to the highest incidence of cancer among all types [2]. Melanoma is a dangerous, uncommon, and fatal form of skin cancer. According to the American Cancer Society, although melanoma skin cancer comprises only 1% of total cases, it exhibits the highest fatality rate [3].

Melanoma begins in cells known as melanocytes. It starts when healthy melanocytes begin to proliferate uncontrollably, resulting in a malignant tumor. It can affect any part of the human body. Moreover, it commonly occurs on sun-exposed regions such as the hands, face, neck, and lips. Melanoma cancers may only be healed if diagnosed early; otherwise, they spread to other body organs and cause terrible death [4]. Meanwhile, nonmelanoma tumors are much easier to treat than melanoma malignancies.

Early detection and accurate classification of skin cancer play a crucial role in improving patient outcomes and reducing mortality rates [5]. In recent years, computer vision and deep learning techniques have shown promising results in skin lesion localization and classification, offering automated and efficient analysis of dermatological images. By leveraging the potential of computer vision algorithms, researchers aim to enhance the speed and accuracy of skin cancer diagnosis, enabling timely interventions and personalized treatment strategies.

The associate editor coordinating the review of this manuscript and approving it for publication was Ines Domingues.

However, skin cancer classification using computer vision still faces several challenges. The research problem lies in developing robust algorithms that can accurately differentiate between benign and malignant skin lesions while considering various factors, such as lesion size, shape, texture, color and handling complex skin lesion patterns. Additionally, the vast diversity in skin types, ethnicities, and imaging conditions further complicates the classification process.

Computer vision-based research in skin cancer detection tends to leverage image analysis and machine learning techniques to improve the accuracy, efficiency, and accessibility of skin cancer diagnosis and management. Within this field, there are several key research areas. Firstly, there is a focus on skin lesion detection and segmentation, accurately identifying and delineating skin lesions within images allowing for subsequent analysis and investigation. Additionally, researchers work on developing efficient methods for feature extraction and representation. Furthermore, the integration of machine learning and deep learning techniques is utilized to create highly accurate classification and risk assessment systems. Another aspect of research involves merging visual analysis with clinical data, including patient history, demographics, and genetic information, to improve diagnostic accuracy and gain a comprehensive understanding of the characteristics of skin lesions [6].

In this paper, various classifiers were tested to classify skin cancer in the images and distinguish between malignant and benign tumors. We used machine learning (ML) methods, such as artificial neural network (ANN), k-nearest neighbor (KNN), support vector machine (SVM), Naïve Bayes (NB), and decision tree (DT), with gray-level co-occurrence matrix (GLCM) feature extractor, which gave lower accuracies compared to deep learning (DL) methods.

AlexNet, VGG-16, VGG-19, EfficientNet-b0, ResNet-18, ResNet-50, ResNet-101, DenseNet-201, Inception-v3, and MobileNet-v2 are ten convolutional neural network (CNN) pretrained DL models that were tested, which provided high performance. We proposed two methods in this paper to get high performance in skin cancer classification. One method is using KNN as classifier when pretrained deep neural networks act as feature extractors (KNN-PDNN) whose accuracies exceeded 99% in some models. The second method is AlexNet optimized by grey wolf optimizer (GWO) (AlexGWO) to enhance the accuracy of AlexNet to 99% by optimizing its parameters using GWO. The implementation of the proposed approaches will allow an early detection and an expedited diagnosis of melanoma. This, in turn, will enable more efficient treatment and reduce the mortality rate associated with the disease. The remainder of the paper is organized as follows. Section II presents the state-of-the-art automated skin cancer binary classification. The materials and methods are shown in Section III. Section IV describes the experimental results. Finally, Section V depicts the discussions followed by the conclusion in section VI.

II. RELATED WORK

This section indicates multiple skin cancer binary classification and detection approaches. It emphasizes recent studies that have employed DL for the same goal.

Tan et al. [7] used particle swarm optimization (PSO) to segment skin lesions. They tested many ways to optimize PSO, including the firefly algorithm, spiral research action, probability distributions, crossover, and mutation. K-means was used to improve lesion segmentation, and CNN was created using the hybrid learning PSO. The categorization method may distinguish between melanoma and nevus lesions.

Kwasigroch et al. [8] proposed employing a CNN with hill climbing for search space to classify skin lesions. This approach resulted in an increase in network size, reducing the computational cost.

Adegun et al. [9] presented an encoder and decoder network with skip links connecting subnetworks. The suggested CNN was employed to segment skin lesions and classify them pixel by pixel.

Song et al. [10] stated that CNNs could segment, identify, and categorize skin lesions. They used a loss function based on the Jaccard distance and the focal loss to regulate the unbalanced datasets.

Lequan et al. [11] presented an extremely deep CNN for melanoma detection. To improve the performance of their method, they employed a fully convolutional residual network with 16 residual blocks in the segmentation process. The suggested approach employed an average of SVM and softmax classifiers for classification. It achieved 85.5% accuracy in melanoma classification with segmentation and 82.8% without segmentation.

DeVries and Ramachandram [12] developed a multi-scale CNN trained on an ImageNet dataset using an Inception-v3 deep neural network. The pretrained Inception-v3 was fine-tuned for skin cancer classification on two resolution scales of input lesion images: coarse and finer scales. The coarse scale was employed to collect morphology and general contextual information about lesions. Meanwhile, the finer scale gathered textual detail of the lesion to distinguish between different forms of skin lesions.

Mahbod et al. [13] suggested a method for extracting deep features from several well-established and pretrained deep CNNs for categorizing skin lesions. Deep-feature generators such as AlexNet, ResNet-18, and VGG-16 were employed, and a multi-class SVM classifier was trained on these produced features. Finally, the results of the classifiers were merged for classification. The suggested method achieved 97.55% and 83.83% area under the curve (AUC) on the ISIC 2017 dataset for seborrheic keratosis and melanoma classification, respectively.

Kalouche et al. [14] suggested a VGG-16 pretrained deep CNN architecture with three fine-tuned layers and five convolutional blocks. VGG-16 models have a 78% accuracy rate in classifying lesion images as melanoma skin cancer.

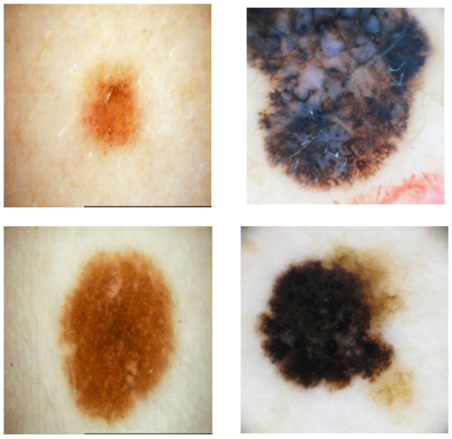

FIGURE 1. Some of ISIC dataset images. (a) benign; and (b) malignant [36].

A deep CNN-based method was proposed to detect the boundaries of skin lesions in photos. The DL model was developed using 1200 photos of normal skin and 400 images of skin lesions. With 86.67% accuracy, the suggested system divided the input photos into two basic classes: normal skin and lesion images.

Various related works of skin cancer binary classification that used ISIC dataset are presented in Table 1.

III. MATERIALS AND METHODS

This paper proposed two methods in skin cancer classification and tested several classification techniques and compared their results on the same dataset. This section demonstrates the used dataset, the preprocessing steps, feature extraction, and the system model with the two proposed methods.

A. DATASET

The used dataset was collected from the ISIC archive. The ISIC Archive is an open-source platform with publicly available dermoscopic images of skin lesions. It includes more than 150,000 total images, of which almost 70,000 have been made public. Images are combined with metadata describing additional characteristics on an image level [36].

We selected 4000 dermoscopic images, from the ISIC archive, divided into 2000 benign and 2000 malignant. The images are in JPEG format. This dataset is randomly split into 80% for training and 20% for testing, which are 3200 and 800 images, respectively. Some of these images from the dataset are shown in Fig. 1.

B. PREPROCESSING

The data were preprocessed as follows. First, because of the different dimensions of ISIC images, it is essential to resize all images to a specific dimension. All images were resized to 250 × 250 pixels. Second, images were filtered using a median filter. Third, the hair of skin that appeared in the images was removed through some morphological

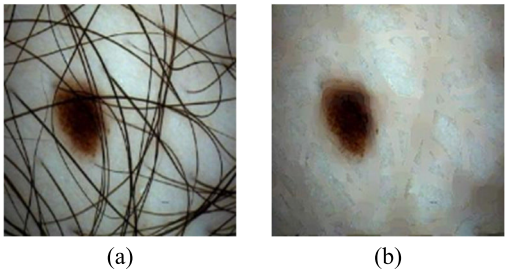

FIGURE 2. Example of hair removal stage. (a) image before hair removal; and (b) image after hair removal.

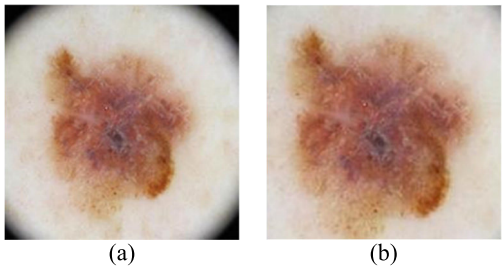

FIGURE 3. Example of image cropping step. (a) image before crop; and (b) image after crop.

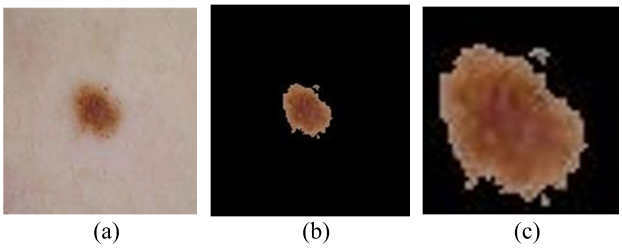

FIGURE 4. Segmentation and cropping around the lesion. (a) Input image; (b) Image after segmentation; and (c) Images after segmentation and surround cropping.

operations not to affect the result of classification as it may be considered as a part of the lesion, as shown in Fig. 2. Fourth, Fig. 3 shows that images were cropped to a proper dimension (150 × 150), as many images have black borders resulting from the microscope. Fifth, we segmented the part of the lesion in all images and discarded the remaining images with black as in Fig. 4(b). Sixth, the black regions around the lesion to get the region of interest, Fig. 4(c); however, this step changed the sizes of the images. Therefore, it must resize all images to the first size (250 × 250). All previous preprocessing steps are presented in Fig. 5. Practically, the segmentation step negatively affected the classification accuracy and reduced the system performance, so we declined this step to get better results.

C. FEATURE EXTRACTION

In this paper, we used GLCM features plus texture features. GLCM is a matrix whose entries represent the number of pairs of pixels with the same brightness level, separated by distance and angle. Angle orientation is determined by four corner directions: 0°, 45°, 90°, and 135°, and the space between pixels is one pixel. The input value of GLCM is a

TABLE 1. Related works of skin cancer binary classification.

Reference	Year	Classification Methods	Dataset	Results (%)
[7]	2020	VGG-16	400 images from ISIC 2017 dataset	Accuracy (73.76)
[8]	2020	VGG-8, VGG-11, VGG-16	ISIC dataset	Accuracy (77)
[9]	2020	CNN	2600 dermoscopy images from ISIC 2017 challenge dataset	Accuracy (95), sensitivity (97), specificity (96)
[10]	2020	Multitask DCNN	13750 images from ISBI 2016 challenge dataset and ISIC 2017 challenge	Accuracy (95.9), sensitivity (83.1), specificity (98.6), dice (95)
[11]	2017	A very deep residual CNN and FCRN	ISIC 2016 database	Accuracy (94.9), sensitivity (91.1), specificity (95.7), Jaccard index (82.9), dice coefficient (89.7)
[12]	2017	Deep multi-scale CNN	ISIC dataset	Accuracy (90.3), AUC (94.3)
[13]	2019	SVM, AlexNet, ResNet-18, VGG-16	ISIC dataset	Average AUC (90.69)
[14]	2016	VGG-16 and CNN	ISIC dataset	Accuracy (78)
[15]	2020	Deep-class CNN	1796 dermoscopy images from ISIC Archive dataset	Accuracy (75), sensitivity (73), specificity (78)
[16]	2021	Region-based CNN with ResNet-152	2742 dermoscopic images from ISIC dataset	Accuracy (90.4), sensitivity (82), specificity (92.5)
[17]	2020	ResNet-50 with deep transfer learning	3600 lesion images from the ISIC dataset	Accuracy (93.5), precision (94), recall (77), F1_score (85)
[18]	2018	Deep CNN	ISIC dataset	Accuracy (80.3), precision (81), AUC (69)
[19]	2019	2-layer CNN with a novel regularizer	ISIC dataset	Accuracy (97.49), AUC (98), sensitivity (94.3), specificity (93.6)
[20]	2017	Hybrid of fully CNN with autoencoder and decoder and RNN	ISIC dataset	Accuracy (98), Jaccard index (93), sensitivity (95), specificity (94)
[21]	2018	6-layers deep CNN	ISIC datasets	Accuracy (77.50)
[22]	2019	CNN	ISIC database	Accuracy (89.5)
[23]	2017	LightNet (DL framework)	ISIC 2016 dataset	Accuracy (81.6), sensitivity (14.9), specificity (98)
[24]	2021	3 different mobile DL models (MobileNet, MobileNet-v2, NASNetMobile)	dataset 2750 skin cancer images from ISIC 2017	Accuracy (NASNetMobile: 82.00), Precision (NASNetMobile: 81.77), F1 Score (MobileNetV2: 81.14), AUC Score (MobileNet: 90.5)
[25]	2021	AttResNet (Attention-based mechanisms and Resnets)	ISIC 2016 and ISIC 2017	Accuracy (93.3)

grayscale image representation matrix. Its output value is a co-occurrence matrix from which features can be extracted based on second-order statistical feature parameters, such as contrast, correlation, homogeneity, and energy.

TABLE 1. *(Continued.)* **Related works of skin cancer binary classification.**

[26]	2022	LR, LDA, KNN, DT, Gaussian NB, VGG-16, Xception, and ResNet-50	ISIC (3297 dermoscopic images)	F-score (88), accuracy (88)
[27]	2021	LeNet, ResNet, EfficientNet, Inception-v3	2019 and 2020 ISIC datasets	Accuracy (91.1)
[28]	2021	VGG, GoogleNet, and ResNet-50	ISIC Archive and SIIM-ISIC 2020	Accuracy (93.7)
[29]	2020	SVM, KNN and NB	1000 dermoscopic images from ISIC 2017	Accuracy (97.8), AUC (94)
[30]	2021	MobileNet-v2, SNN	ISIC	Accuracy (95.27)
[31]	2020	ResNet, AlexNet	ISIC 2019	Accuracy (97.5), F1 score (97.47)
[32]	2022	CNN	1200 dermoscopic images from ISIC 2016, 2017, and 2020	Accuracy (97.5)
[33]	2023	CNN, Xception, ResNet-50, ResNet-101, VGG16, and MC-SVM	ISIC 2018 and ISIC 2019	Accuracy (ISIC 2018: 98.62%, ISIC 2019: 93.47%)
[34]	2022	DenseNet77	ISIC 2017 and ISIC 2018	Accuracy (ISIC 2017: 99.21%, ISIC 2018: 99.51%)
[35]	2022	VGG	ISIC 2016	We can't get them

CNN = Convolutional Neural Network, ISIC = International Skin Imaging Collaboration, SVM = Support Vector Machine, FCRN = Fully Convolutional Residual Network, RNN = Recurrent Neural Network, HAM10000 = Human Against Machine with 10000 training images, PH^2 = Hospital Pedro Hispano, SIIM-ISIC = Society for Imaging Informatics in Medicine, SNN = Spiking Neural Network, NB = Naïve Bayes, LR = Logistic Regression, LDA = Linear Discriminant Analysis, KNN = K-Nearest Neighbor, DT = Decision Tree, DCNN = Deep Convolutional Neural Network, ISBI = IEEE International Symposium on Biomedical Imaging, ResNet = Residual Network, VGG = Visual Geometry Group, MC-SVM = Multi-Class SVM.

Using statistical moments of an image's or region's intensity histogram is one of the easiest methods for defining texture. Using solely histograms in the computation results in texture measurements that contain only information about the distribution of intensities, but not about the relative location of pixels in that texture regarding each other. Using a statistical method, such as a co-occurrence matrix, can offer useful information about the relative location of nearby pixels in a picture.

Given an image I of size NN, the co-occurrence matrix P is defined as follows:

$$P(i,j) = \sum_{x=1}^{N}\sum_{y=1}^{N} \begin{cases} 1, & if\ I(x,y) = i\ and \\ & I\left(x+\Delta_x, y+\Delta_y\right) = j \\ 0, & otherwise \end{cases} \quad (1)$$

The offset (Δ_x, Δ_y) specifies the distance between the pixel-of-interest and its neighbor in this case. The offset (Δ_x, Δ_y) parameterization makes the co-occurrence matrix rotation sensitive. Choosing an offset vector that causes the image's rotation to be less than 180 degrees results in a different co-occurrence matrix for the same (rotated) image. To ensure rotational invariance, create the co-occurrence matrix with a series of offsets sweeping across 180 degrees at the same distance parameter Δ (i.e., [0 Δ] for 0°: P horizontal, [$-\Delta$ Δ] for 45°: P right diagonal, [$-\Delta$ 0] for 90°: P vertical, and [$-\Delta$ $-\Delta$] for 135°: P left diagonal).

Haralick et al. [37] proposed 14 statistical characteristics in 1973. These characteristics are produced by computing the features for each of the co-occurrence matrices created by utilizing the directions 0°, 45°, 90°, and 135°, and then averaging these four values. The distance parameter is represented by the symbol Δ, which might be one or higher. In typically, the distance parameter Δ is set to 1 [38].

In this paper, we extracted 22 features from each image, that were mentioned in [37], [39], and [40]. These 22 features are contrast, correlation, homogeneity, energy, entropy, dissimilarity, inverse difference, autocorrelation, cluster shade, cluster prominence, Maximum probability, Sum of Squares, Sum Average, Sum Variance, Sum Entropy, Difference variance, Difference entropy, Information measures of correlation 1, Information measures of correlation 2, Maximal correlation coefficient, Inverse difference normalized, and Inverse difference moment normalized. We appended four additional features called texture features, which are mean, variance, skewness, and kurtosis to the 22 features to obtain 26 feature parameters for each image.

The following are explanations and calculations for commonly selected GLCM features [38], [41]:

Correlation: The correlation measures the linear relationship of grey levels in neighboring pixels. Correlation is widely utilized in various engineering domains, although it is most employed to quantify displacement, optical flow, and deformation.

$$f_1 = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \frac{(i-\mu_x) \times \left(j-\mu_y\right) \times p(i,j)}{\sigma_x \times \sigma_y} \quad (2)$$

Contrast: The difference in color or hue of any object is known as contrast, and it distinguishes it from other objects.

It may be determined in the actual world by a difference in brightness or hue. It may be calculated mathematically using the following equation.

$$f_2 = \sum_{n=0}^{N_{g-1}} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \right\} \tag{3}$$

Entropy: The amount of information in any image that is required to compress that image exemplifies entropy. It is a numerical measure of uncertainty used to represent the image's texture.

$$f_3 = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p\,(i,j) \times \log\{p(i,j)\} \tag{4}$$

Homogeneity: It uses the value to compute the stiffness of the element distribution. It may be written as follows:

$$f_4 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i,j)}{1+(i-j)^2} \tag{5}$$

Energy: The energy in GLCM is calculated as the sum of square components. Its value spans from [0,1], however for the constant image, it will always be 1. It can be expressed as:

$$f_5 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p\,(i,j)^2, \tag{6}$$

where $p\,(i,j)$ is the (i, j)-th entry in normalized co-occurrence matrix, Ng denotes the dimension of co-occurrence matrix (number of gray levels), μ_x and μ_y are the mean of p_x and p_y, respectively, σ_x and σ_y are the standard deviations of p_x and p_y, respectively, and $p_x(i)$ and $p_y(j)$ are the marginal probabilities.

D. SYSTEM MODEL

In this paper, we proposed two methods: KNN-PDNN and AlexGWO. We also tested different methods to classify dermoscopic images as benign or malignant and compared their results with the proposed ones. These different methods are parts of two approaches, which are, ML and DL, as illustrated in Fig. 6.

1) ML APPROACH

The supervised ML classification algorithms use labeled data to train the classifier. The benefit of supervised over unsupervised is that it may be used for images containing a lot of spectral information. We tested five supervised classification methods: ANN, SVM, KNN, DT, and NB.

a: ANN

ANN is a nonlinear and statistical prediction approach. Its structure is based on the biological structure of the human brain. An ANN consists of three layers of neurons. The first layer is known as the input layer, and its neurons convey data

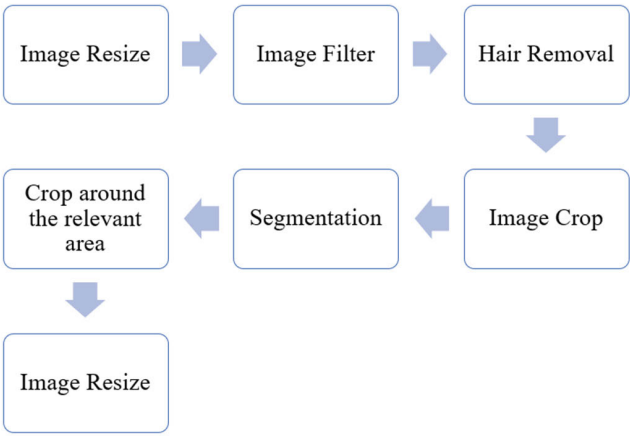

FIGURE 5. Preprocessing steps of the dataset images.

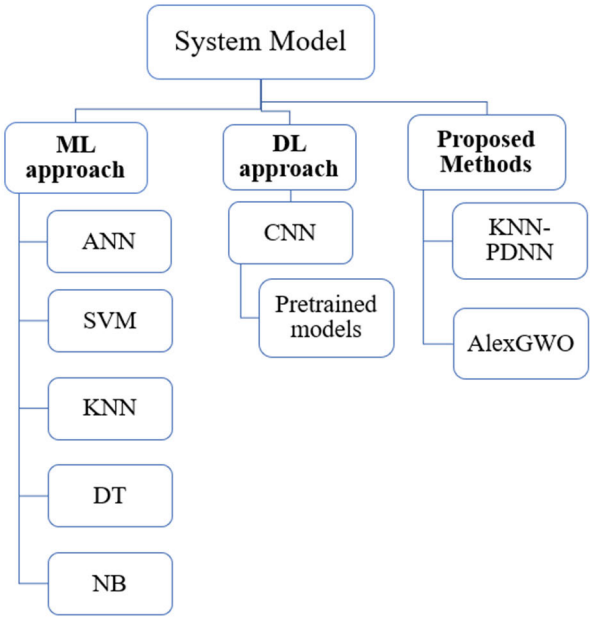

FIGURE 6. System model that was used and tested in this paper.

to the second/intermediate layer of neurons. The intermediate layers are known as hidden layers. A conventional ANN may include many hidden levels. Intermediate neurons transmit information to the third layer of output neurons. Backpropagation (BP) is used to learn the complicated associations/relationships between input and output layers at each layer. It is comparable to a neural network. In computer science, the terms neural network and ANN are now used interchangeably.

ANN is used in skin cancer detection systems to classify retrieved characteristics. After successful training/ classification of the training set, input images are identified as melanoma or nonmelanoma. The number of input images determines the number of hidden layers in an ANN. The input dataset connects the input/first layer of the ANN process with the hidden layer. The dataset can be labeled or unlabeled and processed using a supervised or unsupervised learning technique. A neural network learns weights at each

network connection/link using BP or feed-forward design. Both systems employ a distinct pattern for the underlying dataset. Feed-forward neural networks only transport data in one way. Only data travels from the input to the output [6].

b: SVM

SVMs are supervised learning algorithms that can be used for classification, regression, and outlier identification [42], [43]. A support vector, a hyperplane, separates them as far as feasible based on their class. In this approach, the vector sets the boundary that helps to categorize a new element, such that it is assigned one of two classes based on which portion of the space it belongs to. This algorithm has a set of parameters that allow one to tweak its internal setup and optimize the classification results [44].

c: KNN

The nearest neighbor algorithms work on finding a preset number of training samples that are closest in the distance to the new location and predicting the label of these. The number of samples might be fixed (KNN learning) or variable depending on the local density of the points (based on the neighbor learning radius). Generally, the distance can be any metric measurement, with the standard Euclidean distance being the most frequent [45]. In this paper, we used KNN with ten nearest neighbors.

d: DT

The DT methods are a nonparametric supervised classification technique that can be used for classification and regression learning. The goal is to learn fundamental decision rules generated from data properties and use them to develop a model that predicts the value of a target variable [46].

e: NB

The NB approach is a supervised ML classification method based on a probabilistic technique that employs Bayes' probability theorem. The NB method implies that the occurrences of characteristics are independent. Because of the independence of the occurrence of the extracted characteristics, probability concerns become considerably useful in most domains that deal with occurrences under unpredictability. The extracted feature matrix is trained in the NB classifier so that it can predict whether the test image is normal or malignant [47].

In this study, after data preprocessing, we used k-fold cross-validation and chose k = 5. Then, after converting the images to grayscale, we extracted 26 features from each image using GLCM. Then, we trained 80% of the dataset (3200 images). After that, we tested 20% of the dataset (800 images) and calculated the classification accuracy. The steps of ML methods are shown in Fig. 7.

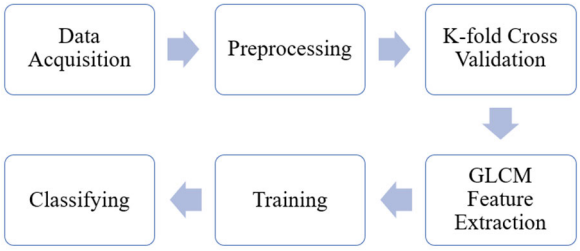

FIGURE 7. Process steps of classification of ML approach.

2) DL APPROACH

DL is a subset of machine learning that is essentially a three- or more-layered neural network. These neural networks try to mimic the human brain function by allowing it to "learn" from vast volumes of data. While a single-layer neural network may still produce approximate predictions, more hidden layers can assist optimize and tune for accuracy. We tested CNN and some CNN pretrained models.

a: CNN

CNN is a form of deep neural network that is widely used in computer vision. It is used to classify images, assemble a set of input images, and recognize images. CNN is an excellent technique for collecting and learning global and local data by combining simple characteristics, such as curves and edges, to construct complex features (e.g., forms and corners) [48]. The hidden layers of CNN consist of convolutional, nonlinear pooling, and fully connected layers [49]. CNN can have many convolutional layers followed by several fully connected layers. Convolutional, pooling, and full-connected layers are the three basic types of layers used in CNN [50].

CNN is widely used in image classification because of convolution techniques' efficacy in shape extraction. The following steps are performed when we use CNN architecture. We loaded the dataset after preprocessing and split them into 80% for training and 20% for testing. Then, we used image data augmentation to resize the original images as the size of the input layer of the CNN (224×224). Then, we defined the layers: five convolution layers with 32, 32, 64, 128, and 128 filters in each, one fully connected layer, a dropout layer before the fully connected layer to avoid overfitting, five maximum pooling layers, five ReLU layers, one softmax layer, and five batch normalization layers each before each ReLU layer to improve the speed and stability of the neural network. Next, we created a set of options for training a network. After that, the network was trained and classified, and its accuracy was computed.

b: CNN PRETRAINED MODELS

A stored network previously trained on a big dataset, generally on a large-scale image-classification job, is referred to as a pretrained model. In this paper, we used ten different DL models: AlexNet, ResNet-18, ResNet-50, ResNet-101, DenseNet-201, VGG-16, VGG-19, EfficientNet-b0, MobileNet-v2, and Inception-v3, which discussed as follows.

i) ALEXNET

AlexNet is a CNN model that primarily influences DL applications to computer vision. It won the 2012 ImageNet LSVRC-2012 competition with a significant margin (15.3% botch rates versus 26.2% blunder rates in the runner-up, which is VGG-16). The organization's configuration was like the LeNet of Yann LeCun et al. but more profound, with more channels per layer and stacked convolutional layers. Convolutions, maximum pooling, dropout, information growth, ReLU initiations, and stochastic gradient descent with force were crucial. It adds ReLU initiations after each convolutional and totally related layer. Furthermore, instead of regularization, dropout is used to cope with overfitting [51].

ii) RESNET-18, 50, AND 101

The ResNet-50 model won the ILSRVC-2015 competition, with a 3.57% error rate and an input image size of 224 by 224 pixels. ResNet is a well-known DL model published by Shaoqing Ren, Kaiming He, Jian Sun, and Xiangyu Zhang. The ResNet-18 model has 18 layers, whereas the ResNet-50 model has 50 layers, each with two or three convolutional layers. ResNet-101 is a DL model with 101 layers [52].

iii) DENSENET-201

Huang et al. [53] presented the DenseNet-201 design, one of the most current dense network variations. Each layer in DenseNet receives more data from previous layers and delivers its characteristic to each subsequent layer. The term ''link'' is used. Each layer receives aggregate data from the previous layers. The DenseNet-201 model also includes a pooling layer and bottleneck development. Here, the mistake may be passed on more directly to previous tiers. Because previous layers can get tight supervision from the last order layer, this is more thorough management, which is the advantage of the model [51]. In this paper, we used DenseNet-201 with 201 layers.

iv) VGG-16 AND 19

VGG is distinguished by its simplicity. It consists of five blocks of 3×3 convolutional layers stacked on each other. Maximum pooling of 2×2 kernels and a stride of two reduce the volume size. It is followed by two completely connected layers with ReLU activation functions, each with 4096 nodes. The last layer has 1000 nodes activated by softmax [54]. VGG-19 has around 143 million parameters. The Vgg-16 model has 16 layers, whereas the Vgg-19 model has 19 layers.

v) EFFICIENTNET

EfficientNet is a CNN design that scales up by combining coefficients on width, height, and resolution, often known as compound scaling [55]. Several studies concluded that EfficientNet might surpass other state-of-the-art methods in terms of accuracy and efficiency [56]. A feature map is applied to each layer for width scaling. Depth scaling increased the number of layers in the network design. Additionally, resolution scaling raised the resolution of the supplied image [57]. The EfficientNet architecture is divided into eight variations: EfficientNet-B0 to EfficientNet-B7. This architecture is built on EfficientNet-B0, which adapts the inverted bottleneck residual block MobileNetV2, commonly known as MBConv. This residual block is employed to improve network efficiency [58].

vi) MOBILENET-V2

The MobileNet-v2 model inherits MobileNet's aims and improves on them. It has fewer layers than MobileNet and requires fewer processing resources. The number of layers in MobileNet-v2 is the same as in MobileNet; however, the number of parameters is reduced to 3.5 million and the size is reduced to 14 MB. MobileNet-v2 improves MobileNet by including shortcut connections, inverted residual blocks, and bottleneck blocks. The use of inverted residual bottleneck layers provides for a more memory-efficient design. This model also performs well in item identification and semantic segmentation [58].

vii) INCEPTION-V3

The Inception-v3 [59] model enhanced the GoogLeNet architecture [60]. The primary concept behind this network is to make the procedure easier and more efficient. The inception module performs the function of a multilevel feature extractor. It computes 1×1, 3×3, and 5×5 convolutions inside the same network module. The outputs of these filters are then layered on top of one another and supplied into the next layer of the network [61].

In this work, pretrained deep networks are tested as classifiers with transfer learning (TL).

TL is an ML strategy that reuses a model produced for one task in another. It is typically used when there is insufficient training data. However, data augmentation can help to overcome the data problem. We require transfer learning because melanoma and benign lesions are very similar, and it takes a long time to recognize and classify them. Furthermore, transfer learning is more efficient in categorizing related lesions, making it the preferred method. Transfer learning networks are trained on huge datasets, and their model weights are frozen before changing the last few layers for a different dataset [61].

DL applications frequently employ TL. A pretrained network can be used as a starting point for learning a new task. TL is considerably faster and easier than training a network from scratch with randomly initialized weights.

We began with loading preprocessing and splitting data. Then, we loaded a pretrained network. The network's convolutional layers retrieved image features that the final learnable and final classification layers used to classify the input image. The classification layer specifies the networ's output classes. We replaced the classification layer with a new one that did not include class labels. We set the learning rates in earlier layers of the network to zero to freeze their weights.

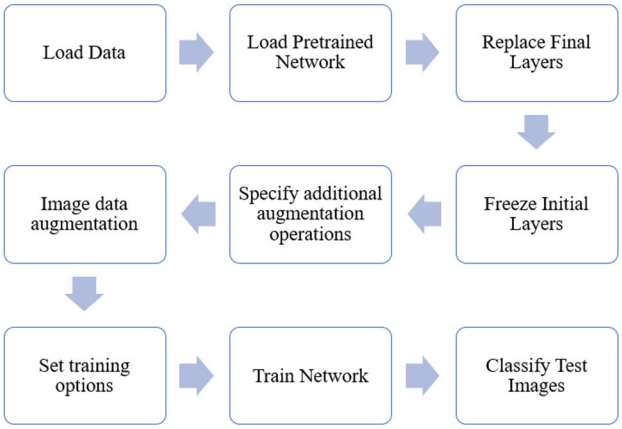

FIGURE 8. Process steps of classification of pretrained models with TL.

The network did not change the parameters of the frozen layers during training. Freezing the weights of several early layers can considerably accelerate network training since the gradients of the frozen layers do not need to be computed.

The network requires 224 × 224 input images; however, the images in the image data store have different sizes. To automatically resize the training images, we used an augmented image data store. We specified additional augmentation operations on the training images, including randomly flipping them along the vertical axis, randomly translating them up to 30 pixels, and scaling them up to 10% horizontally and vertically. Data augmentation prevents the network from overfitting and remembering the exact characteristics of the training images. The training options were then specified and set as the number of epochs to be trained for. It does not need to train for as many epochs when using transfer learning. We used the fine-tuned network to classify the testing images and determine the classification accuracy. Fig. 8 shows the steps of pretrained models as classifiers with TL.

3) PROPOSED METHODS

In this paper, we proposed two methods to detect and classify skin cancer disease: KNN-PDNN and AlexGWO.

a: KNN-PDNN

In this method, the classifier KNN was used with the ten used pretrained models, that were mentioned previously, as feature extractors. We tested some of ML methods, e.g., SVM, KNN, DT, and NB, to find the best performance. KNN (with the number of nearest neighbors equals ten) gave high accuracies compared to other methods. Thus, we used KNN as classifier when these models were used as feature extractors. After loading preprocessing data, splitting it, and loading a pretrained network, the network requires input images with a size of 227 × 227, although the images in the image data stores have different sizes. We developed augmented image data stores and provided the required image size to automatically resize the training and test images before they were submitted to the network.

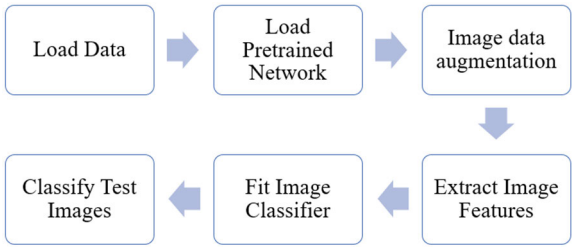

FIGURE 9. Process steps of classification of proposed KNN-PDNN.

We loaded the data by splitting it into 80% training and 20% test data, then loaded one of the pretrained network. Deeper layers of a pretrained model have higher-level features that are built on the lower-level elements of prior levels. We employed the activation function on the global pooling layer, which changes from model to model, to acquire the feature representations of the training and testing images at the end of the network. The input features are pooled over all spatial locations in the global pooling layer. The models were then used to extract train and test features. We employed features retrieved from the training images as predictor variables, fitted KNN, classified the test images using the trained KNN model and the features collected from the test images, and calculated the classification accuracy of the test set, as illustrated in Fig. 9.

b: ALEXGWO

In this study, we selected the nature inspired GWO method [62] for various reasons, including its few tuning parameters, quick convergence, and capacity to handle optimization difficulties. The GWO algorithm is based on the hunting concept. Wolves are part of a group that includes several grey wolves that help in hunting. Wolves in a pack are classified based on their ability to lead. A pack of wolves is divided into four types: alpha (α), beta (β), delta (δ), and omega (ω). The leader of the hunting process is the group's decision-maker.

The dominance of the remaining wolves decreases gradually in the following order: β, δ,and ω. Such wolves participate in hunting and pass on their improved positions to their superiors. The earliest grey wolves look for and contain prey in this phase, which is the course of the chase. α, β,and δ have more knowledge about prospective prey areas for mathematical simulations of hunting behavior [63].

Seyedali Mirjalili introduced GWO in 2014 by emulating the social behavior, leadership structure, and hunting in the collective property of grey wolves [62]. Grey wolves usually live in packs in the wild, with group sizes ranging from 5 to 12. They maintain a rigid social dominance hierarchy. In the highest level of the hierarchy, the most dominant male or female wolves are portrayed as α, and they are primarily decision-makers for the wolf pack's feeding, sleeping, hunting, and habitat. The α wolf is followed by all the other wolves. The β wolves are the next level wolves in the hierarchy; they obey the α and govern the lower-level wolves. The δ wolves assist the α and β wolves in hunting and seeking

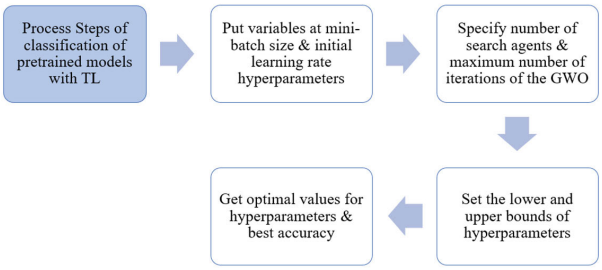

FIGURE 10. Process steps of proposed AlexGWO.

prey in the following category. They patrol the region, alert the other wolves to any danger, and care for the injured and weak wolves.

Furthermore, ω wolves are the lowest category and must obey all other wolves' orders. Wolves' hunting success is primarily based on their social hierarchy. Grey wolf social behavior may be mathematically described by selecting the prey location as the ideal solution and describing the wolf position as the solution in the search space. Moreover, α wolves are the greatest solution because of their proximity to the prey. According to their social structure, the β and δ wolves are the next best solutions. Meanwhile, ω wolves in search space adjust their position based on the positions of α, β, and δ wolves. GWO's main phases are prey encirclement, hunting, fighting, and searching. Prey encircling refers to how wolves encircle their prey when hunting. The prey hunting procedure is led by the α wolves. The β and δ are also involved in this process. It is anticipated that these three wolves will be informed about the likely prey zone. This aids in selecting the three finest search agents, who then assist in updating the locations of other wolves. The seeking or exploring for an optimal solution is modeled after wolf searching behavior. Wolves divide when looking for prey and converge when they locate it [64].

AlexNet was chosen with GWO because AlexNet has the least depth of any pretrained deep models; thus, it takes less time to execute. The same steps of pretrained models as classifiers were performed except that we put variables at the mini-batch size and the initial learning rate hyperparameters in the training options to be optimized by GWO, as shown in Fig. 10. We also specified the number of search agents and the maximum number of iterations of GWO. Then, we set the lower and upper bounds for the required parameters. In details, we initialized alpha, beta, and delta positions, initialized the positions of search agents, returned the search agents that go beyond the boundaries of the search space, calculated objective function for each search agent, updated alpha, beta, and delta, and updated the position of search agents including omegas. A diagram of the abbreviated steps is presented in Fig. 10.

IV. EXPERIMENTAL RESULTS

Several tests were performed in this section to evaluate the performance of the proposed approach and several state-of-the-art classification algorithms on the ISIC archive dataset.

A. SYSTEM IMPLEMENTATION

The implemented frameworks have been tested and evaluated using the following software and hardware configurations:

- Operating system: Windows 10 Pro.
- Compiler: MATLAB R2020b.
- Processor: Intel (R) Core (TM) i7-9750H CPU @ 2.60GHz 2.59 GHz.
- Installed RAM: 16.0 GB (15.9 GB usable).
- System type: 64-bit operating system, x64-based processor.

B. EVALUATION METRICS

We employ five popular measures to objectively evaluate the effectiveness of the proposed melanoma detection method: precision, sensitivity/recall, specificity, F1 score, and accuracy:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{7}$$

$$\text{Sensitivity/Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{8}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{9}$$

$$\text{F1score} = \frac{2 \times \text{TP}}{2 \times \text{TP} + \text{FP} + \text{FN}} \tag{10}$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \tag{11}$$

where TP (True Positive) is the number of Malignant cases correctly classified, TN (True Negative) is the number of Benign cases correctly classified, FP (False Positive) is the number of Benign cases detected as Malignant, and FN (False Negative) is the number of Malignant cases detected as Benign.

C. RESULTS AND DISCUSSION

This section presents the results and the used parameters of the tested classification approaches and the two proposed methods:

In ML approach, we tested ANN, KNN (with the number of nearest neighbors equals ten), NB, SVM, and DT with GLCM feature extraction before classification and extracted 26 features of images. ANN has 26 inputs (features), one hidden layer with 52 neurons, and one output (benign or malignant) with 1000 epochs. We split data into 80% train and 20% test by k-fold cross-validation (k = 5). Classification accuracies of ANN, KNN, NB, SVM, and DT models are 81.25%, 72.35%, 70.45%, 64.92%, and 77.25%, respectively. The results are presented in Fig. 11.

In DL approach, we employed CNN with five convolutional layers, and one fully connected layer. The number of filters in each convolutional layer is 32, 32, 64, 128, and 128. Its training options were set as the mini-batch size is 32, and maximum epochs to 5. We also tested 10 and

TABLE 2. Performance of pretrained deep models. the best values are shown in bold and underline.

Pretrained models	Precision (%)	Sensitivity/Recall (%)	Specificity (%)	F1 Score (%)	Accuracy (%)
Densenet-201	**98.0344**	99.75	**98**	**98.8848**	**98.875**
ResNet-101	96.6184	**100**	96.5	98.2801	98.25
AlexNet	96.6019	99.5	96.5	98.0296	98
ResNet-18	95.6731	99.5	95.5	97.549	97.5
ResNet-50	93.662	99.75	93.25	96.6102	96.5
MobileNet-v2	97.1503	93.75	97.25	95.4198	95.5
Inception-v3	90.4762	99.75	89.5	94.887	94.625

TABLE 3. Performance of the proposed KNN-PDNN. the best values are shown in bold and underline.

Pretrained models	Precision (%)	Sensitivity/Recall (%)	Specificity (%)	F1 Score (%)	Accuracy (%)
VGG-16	**99.7481**	99	**99.75**	**99.3726**	**99.375**
EfficientNet-B0	99.0025	**99.25**	99	99.1261	99.125
VGG-19	98.9975	98.75	99	98.8736	98.875
ResNet-50	99.2424	98.25	99.25	98.7437	98.75
ResNet-101	98.7469	98.5	98.75	98.6233	98.625
ResNet-18	98.9924	98.25	99	98.6198	98.625
Densenet-201	98.7245	96.75	98.75	97.7273	97.75
Inception-v3	96.5261	97.25	96.5	96.8867	96.875
AlexNet	99.2	93	99.25	96	96.125
MobileNet-v2	98.3651	90.25	98.5	94.133	94.375

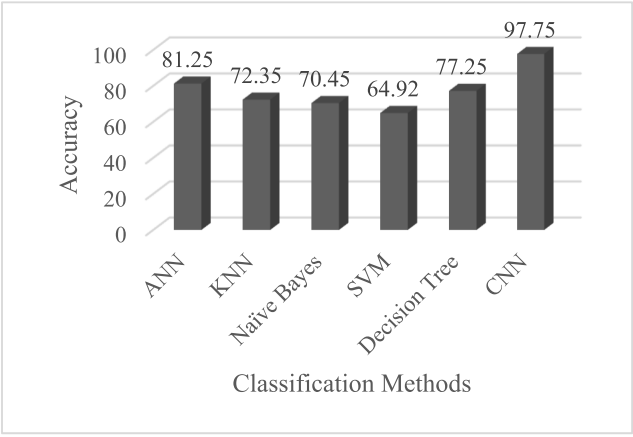

FIGURE 11. Accuracy comparison chart among ML methods and CNN.

15 epochs but, in these cases, we suffered from complexity in processing time besides it provided only small differences in performance. The initial learning rate is 1×10^{-4} and its accuracy is 97.75%. Fig. 11 shows the comparison between ML methods and CNN based on their accuracies. As shown in Fig. 11, there is a considerable difference in accuracies between CNN and ML methods.

We tested pretrained models: VGG-16, VGG-19, EfficientNet-B0, AlexNet, ResNet-18, ResNet-50, ResNet-101, DenseNet-201, Inception-v3, and MobileNet-v2. These models were tested as classifiers with transfer learning. Their training options were chosen as follows: maximum epochs are 5, mini-batch size is 32, and initial learning rate is 1×10^{-4}. The confusion matrix of the highest accuracy, Densenet-201, is shown in Fig. 12. The results of the pretrained models are shown in Table 2.

In the first proposed method, KNN-PDNN, pretrained models were deployed as feature extractors when we used KNN as a classifier. When the number of nearest neighbors equals ten, KNN outperformed other approaches, whose accuracy exceeds 99% in some models. The mini-batch size was 32. Table 3 shows that VGG-16 has the best accuracy. The confusion matrix of VGG-16 is shown in Fig. 13.

In the second proposed method, AlexGWO, the number of search agents in GWO was set to 10, and the maximum number of iterations was 50. We set the lower and upper bounds for the mini-batch size to be 2 and 128, respectively. The lower and upper bounds of the initial learning rate are 0.0001 and 0.000001, respectively. The best parameters that give the best accuracy are a mini-batch size of 40 and initial learning rate of 9×10^{-4}, which provide the best accuracy for

TABLE 4. Performance of the proposed ALEXGWO.

Method	Precision (%)	Sensitivity/Recall (%)	Specificity (%)	F1 Score (%)	Accuracy (%)
AlexGWO	99.47	100	99.5	98.63	99

FIGURE 12. Confusion matrix of Densenet-201.

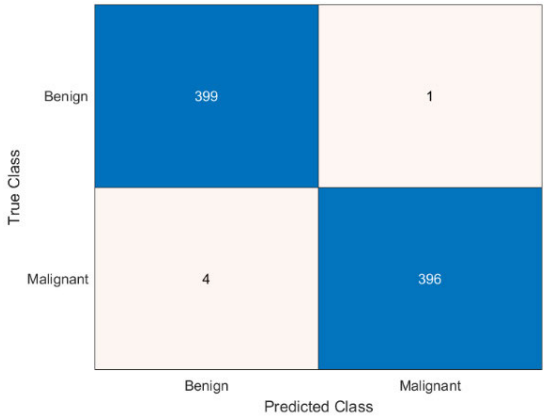

FIGURE 13. Confusion matrix of VGG-16.

AlexNet (99%) compared to Alexnet as feature extractor with KNN classifier (96.125%) and Alexnet as classifier (98%). Performance of proposed AlexGWO is shown in Table 4.

Table 5 compares the accuracy of the proposed methods with those of the existing DL methods on the same dataset. It shows that the two proposed methods achieve the best accuracy.

V. DISCUSSIONS

We proposed two classification approaches based on deep learning networks to automatically detect skin cancer in dermoscopy images, while also conducting a wide range of tests to assess the performance of the proposed methods.

The performances of the two proposed models, KNN-PDNN and AlexGWO, are compared to most of ML and DL approaches. By comparing the accuracy values acquired from other evaluated approaches, we can see that our proposed methods outperform other tested methods.

TABLE 5. Comparison of accuracies between existing dl works and the proposed methods on isic dataset. the highest accuracy is shown in bold and underline.

Classification Method	Accuracy (%)
VGG-16 [7]	73.76
VGG-8, VGG-11, VGG-16 [8]	77
CNN [9]	95
Multitask DCNN [10]	95.9
A very deep residual CNN and FCRN [11]	94.9
Deep multi-scale CNN [12]	90.3
VGG-16 and CNN [14]	78
Deep-class CNN [15]	75
Region-based CNN with ResNet-152 [16]	90.4
ResNet-50 with deep transfer learning [17]	93.5
Deep CNN [18]	80.3
CNN [22]	89.5
LightNet (DL framework) [23]	81.6
MobileNet, MobileNet-v2, NASNetMobile [24]	82
VGG-16, Xception, and ResNet-50 [26]	88
LeNet, ResNet, EfficientNet, Inception-v3 [27]	91.1
VGG, GoogleNet, and ResNet-50 [28]	93.7
MobileNet-v2 [30]	95.27
ResNet, AlexNet [31]	97.5
CNN [32]	97.5
Proposed KNN-PDNN	**99.375**
Proposed AlexGWO	**99**

The proposed models can be easily accessible via a Web-based platform or even as an API to aid dermatologists in detecting possible lesion risk in a timely manner; this is surely another subject that we will investigate.

The next goal is to analyze the dynamic evolution of the skin lesion over a short period of time to improve classification and prediction. As a result, we aim to expand the taxonomy to include different forms of skin lesions that may be associated with a cancer diagnosis.

More dermoscopy images can be added to the training dataset in the future to improve and increase efficiency. To make the model more diversified in nature, datasets from various age groups and categories may be incorporated. To improve prediction efficiency, picture metadata can be supplied. Based on the patient's medical history and other personal information, a personalized system can be created.

The objective of this research field has always been to provide medical-grade apps to aid in diagnosis. With the increase of skin cancer as a worldwide health concern, and melanoma being the deadliest variety, it is anticipated that the information presented in this study will help to advance these technologies.

VI. CONCLUSION

The goal of medical disease classification research is to help patients and clinicians. Doctors can make better judgments when using computer-based solutions. An accurate approach promotes faith in an illness diagnosis. When identifying skin

cancer, it is challenging to distinguish between benign and malignant lesions. Therefore, in this study, we examined the effectiveness of ML and DL models in interpreting dermoscopic images of skin lesions. Although, current ML and DL experiments have piqued the public's curiosity and shown great potential, such image processing has long been challenging and error prone. Based on only dermoscopic images, we developed ML and DL models to examine whether a tumor is malignant or benign. We observed that the basic ML algorithms were less accurate on the tested 4000 dermoscopic images from the ISIC dataset. When images were examined using CNN and pretrained deep networks, there were apparent and discernible differences in accuracy compared to previous methodologies. We achieved accuracy that exceeded 99% in the first proposed method, KNN-PDNN, and 99% in the second proposed method, AlexGWO. Furthermore, to improve the model's performance, it is better to employ DL techniques to construct a model for categorizing benign and malignant skin tumors. DL reduces error while improving accuracy.

REFERENCES

[1] R. Ashraf, S. Afzal, A. U. Rehman, S. Gul, J. Baber, M. Bakhtyar, I. Mehmood, O. Song, and M. Maqsood, "Region-of-interest based transfer learning assisted framework for skin cancer detection," *IEEE Access*, vol. 8, pp. 147858–147871, 2020.

[2] A. L. Byrd, Y. Belkaid, and J. A. Segre, "The human skin microbiome," *Nature Rev. Microbiol.*, vol. 16, pp. 143–155, Jan. 2018.

[3] American Cancer Society. *About Melanoma Skin Cancer*. Accessed: Dec. 10, 2022. [Online]. Available: https://www.cancer.org/content/dam/CRC/PDF/Public/8823.00.pdf

[4] M. Q. Khan, A. Hussain, S. U. Rehman, U. Khan, M. Maqsood, K. Mehmood, and M. A. Khan, "Classification of melanoma and nevus in digital images for diagnosis of skin cancer," *IEEE Access*, vol. 7, pp. 90132–90144, 2019.

[5] P. S. Staff, *What's the Difference Between Melanoma and Non-Melanoma Skin Cancer?* Accessed: Dec. 10, 2022. [Online]. Available: https://www.premiersurgical.com/01/whats-the-difference-between-melanoma-and-non-melanoma-skin-cancer/

[6] M. Dildar, S. Akram, M. Irfan, H. U. Khan, M. Ramzan, A. R. Mahmood, S. A. Alsaiari, A. H. M. Saeed, M. O. Alraddadi, and M. H. Mahnashi, "Skin cancer detection: A review using deep learning techniques," *Int. J. Environ. Res. Public Health*, vol. 18, p. 5479, May 2021.

[7] T. Y. Tan, L. Zhang, and C. P. Lim, "Adaptive melanoma diagnosis using evolving clustering, ensemble and deep neural networks," *Knowl.-Based Syst.*, vol. 187, Jan. 2020, Art. no. 104807.

[8] A. Kwasigroch, M. Grochowski, and A. Mikolajczyk, "Neural architecture search for skin lesion classification," *IEEE Access*, vol. 8, pp. 9061–9071, 2020.

[9] A. A. Adegun and S. Viriri, "Deep learning-based system for automatic melanoma detection," *IEEE Access*, vol. 8, pp. 7160–7172, 2020.

[10] L. Song, J. Lin, Z. J. Wang, and H. Wang, "An end-to-end multi-task deep learning framework for skin lesion analysis," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 10, pp. 2912–2921, Oct. 2020.

[11] L. Yu, H. Chen, Q. Dou, J. Qin, and P. Heng, "Automated melanoma recognition in dermoscopy images via very deep residual networks," *IEEE Trans. Med. Imag.*, vol. 36, no. 4, pp. 994–1004, Apr. 2017.

[12] T. DeVries and D. Ramachandram, "Skin lesion classification using deep multi-scale convolutional neural networks," 2017, *arXiv:1703.01402*.

[13] A. Mahbod, G. Schaefer, C. Wang, R. Ecker, and I. Ellinge, "Skin lesion classification using hybrid deep neural networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 1229–1233.

[14] S. Kalouche, A. Ng, and J. Duchi, "Vision-based classification of skin cancer using deep learning," Stanford Univ., Stanford, CA, USA, Tech. Rep., 2016.

[15] S. Nasiri, J. Helsper, M. Jung, and M. Fathi, "DePicT melanoma deepCLASS: A deep convolutional neural networks approach to classify skin lesion images," *BMC Bioinf.*, vol. 21, no. 2, pp. 1–13, Mar. 2020.

[16] M. F. Jojoa Acosta, L. Y. Caballero Tovar, M. B. Garcia-Zapirain, and W. S. Percybrooks, "Melanoma diagnosis using deep learning techniques on dermatoscopic images," *BMC Med. Imag.*, vol. 21, no. 1, pp. 1–11, Jan. 2021.

[17] A. Sagar and J. Dheeba, "Convolutional neural networks for classifying melanoma images," bioRxiv, May 2020.

[18] V. Singh and I. Nwogu, "Analyzing skin lesions in dermoscopy images using convolutional neural networks," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2018, pp. 4035–4040.

[19] M. A. Albahar, "Skin lesion classification using convolutional neural network with novel regularizer," *IEEE Access*, vol. 7, pp. 38306–38313, 2019.

[20] M. Attia, M. Hossny, S. Nahavandi, and A. Yazdabadi, "Skin melanoma segmentation using recurrent and convolutional neural networks," in *Proc. IEEE 14th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2017, pp. 292–296.

[21] M. S. S. Mahecha, O. J. S. Parra, and J. B. Velandia, "Design of a system for melanoma detection through the processing of clinical images using artificial neural networks," in *Challenges and Opportunities in the Digital Era* (Lecture Notes in Computer Science). Berlin, Germany: Springer, 2018, pp. 605–616.

[22] D. C. Malo, Md. M. Rahman, J. Mahbub, and M. M. Khan, "Skin cancer detection using convolutional neural network," in *Proc. IEEE 12th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Jan. 2022, pp. 0169–0176.

[23] A. A. Ali and H. Al-Marzouqi, "Melanoma detection using regular convolutional neural networks," in *Proc. Int. Conf. Electr. Comput. Technol. Appl. (ICECTA)*, Nov. 2017, pp. 1–5.

[24] A. Yilmaz, M. Kalebasi, Y. Samoylenko, M. E. Guvenilir, and H. Uvet, "Benchmarking of lightweight deep learning architectures for skin cancer classification using ISIC 2017 dataset," 2021, *arXiv:2110.12270*.

[25] S. Y. Boulahia, M. A. Benatia, and A. Bouzar, "Att2ResNet: A deep attention-based approach for melanoma skin cancer classification," *Int. J. Imag. Syst. Technol.*, vol. 32, no. 2, pp. 476–489, Dec. 2021.

[26] S. Bechelli and J. Delhommelle, "Machine learning and deep learning algorithms for skin cancer classification from dermoscopic images," *Bioengineering*, vol. 9, no. 3, p. 97, Feb. 2022.

[27] Q. Deng, J. C. C. Beltran, and D. Lee, "Assessment of segmentation impact on melanoma classification using convolutional neural networks," *J. Comput. Sci. Eng.*, vol. 15, no. 3, pp. 115–124, Sep. 2021.

[28] Z. M. Arkah, D. S. Al-Dulaimi, and A. R. Khekan, "Big transfer learning for automated skin cancer classification," *Indonesian J. Electr. Eng. Comput. Sci.*, vol. 23, no. 3, p. 1611, Sep. 2021.

[29] M. Vidya and M. V. Karki, "Skin cancer detection using machine learning techniques," in *Proc. IEEE Int. Conf. Electron., Comput. Commun. Technol. (CONECCT)*, Jul. 2020, pp. 1–5.

[30] M. Togaçar, Z. Cömert, and B. Ergen, "Intelligent skin cancer detection applying autoencoder, MobileNetV2 and spiking neural networks," *Chaos, Solitons Fractals*, vol. 144, Mar. 2021, Art. no. 110714.

[31] L. Ichim and D. Popescu, "Melanoma detection using an objective system based on multiple connected neural networks," *IEEE Access*, vol. 8, pp. 179189–179202, 2020.

[32] T. Vaiyapuri, P. Balaji, H. Alaskar, and Z. Sbai, "Computational intelligence-based melanoma detection and classification using dermoscopic images," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–12, May 2022.

[33] S. Maqsood and R. Damasevicius, "Multiclass skin lesion localization and classification using deep learning based features fusion and selection framework for smart healthcare," *Neural Netw.*, vol. 160, pp. 238–258, Mar. 2023.

[34] M. Nawaz, T. Nazir, M. Masood, F. Ali, M. A. Khan, U. Tariq, N. Sahar, and R. Damasevicius, "Melanoma segmentation: A framework of improved DenseNet77 and UNET convolutional neural network," *Int. J. Imag. Syst. Technol.*, vol. 32, no. 6, pp. 2137–2153, May 2022.

[35] V. Rajinikanth, S. Kadry, R. Damasevicius, D. Sankaran, M. A. Mohammed, and S. Chander, "Skin melanoma segmentation using VGG-UNet with Adam/SGD optimizer: A study," in *Proc. 3rd Int. Conf. Intell. Comput. Instrum. Control Technol. (ICICICT)*, Aug. 2022, pp. 982–986.

[36] ISIC-Archive. *Content and Layout of the Archive*. Accessed: Dec. 10, 2022. [Online]. Available: https://www.isic-archive.com/#!/topWithHeader/tightContentTop/about/isicArchiveContent

[37] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973.

[38] A. Eleyan and H. Demirel, "Co-occurrence matrix and its statistical features as a new approach for face recognition," *Turkish J. Electr. Eng. Comput. Sci.*, pp. 97–107, Jan. 2011.
[39] L.-K. Soh and C. Tsatsoulis, "Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 2, pp. 780–795, Mar. 1999.
[40] D. A. Clausi, "An analysis of co-occurrence texture statistics as a function of grey level quantization," *Can. J. Remote Sens.*, vol. 28, no. 1, pp. 45–62, Jan. 2002.
[41] Z. Abbas, M. Rehman, S. Najam, and S. M. D. Rizvi, "An efficient gray-level co-occurrence matrix (GLCM) based approach towards classification of skin lesion," in *Proc. Amity Int. Conf. Artif. Intell. (AICAI)*, Feb. 2019, pp. 317–320.
[42] D. H. S. Prashantha and B. R. Shetty, "Comparison of image classification techniques and its applications," *ASIAN J. Converg. Technol.*, vol. 7, no. 1, pp. 58–62, Apr. 2021.
[43] V. Vapnik, *The Nature of Statistical Learning Theory*. Berlin, Germany: Springer, 2013.
[44] I. Gaytán-Campos, W. Morales-Castro, B. Priego-Sánchez, E. Fitz-Rodríguez, and R. Guzmán-Cabrera, "Automatic classification of images with skin cancer using artificial intelligence," *Computación y Sistemas*, vol. 26, no. 1, pp. 325–336, Mar. 2022.
[45] B. Tu, J. Wang, X. Kang, G. Zhang, X. Ou, and L. Guo, "KNN-based representation of superpixels for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 11, pp. 4032–4047, Nov. 2018.
[46] H. Liu, M. Cocea, and W. Ding, "Decision tree learning based feature evaluation and selection for image classification," in *Proc. Int. Conf. Mach. Learn. Cybern. (ICMLC)*, vol. 2, Jul. 2017, pp. 569–574.
[47] H. T. Zaw, N. Maneerat, and K. Y. Win, "Brain tumor detection based on Naïve Bayes classification," in *Proc. 5th Int. Conf. Eng., Appl. Sci. Technol. (ICEAST)*, Jul. 2019, pp. 1–4.
[48] M. U. Rehman, S. H. Khan, S. M. D. Rizvi, Z. Abbas, and A. Zafar, "Classification of skin lesion by interference of segmentation and convolotion neural network," in *Proc. 2nd Int. Conf. Eng. Innov. (ICEI)*, Jul. 2018, pp. 81–85.
[49] A. W. Harley, "An interactive node-link visualization of convolutional neural networks," in *Proc. Adv. Vis. Comput.* Cham, Switzerland: Springer, 2015, pp. 867–877.
[50] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," 2013, *arXiv:1311.2901*.
[51] S. Suganyadevi and V. Seethalakshmi, "DarkCVNet: Optimized pneumonia and COVID 19 detection using CXR images," KPRIET, KPR Inst. Eng. Technol., Coimbatore, India, Tech. Rep., Jul. 2022.
[52] K. Vyshnavi, A. Kintali, K. Sravan, V. Kothuru, and V. Gowtham, "Performance analysis of pneumonia detection using resnets technique," *Int. J. Sci. Eng.*, vol. 6, no. 9, pp. 5–12, 2021, doi: 10.51397/OAIJSE09.2021.0002.
[53] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2261–2269.
[54] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[55] A. M. K. Izzaty, T. W. Cenggoro, G. N. Elwirehardja, and B. Pardamean, "Multiclass classification of histology on colorectal cancer using deep learning," *Commun. Math. Biol. Neurosci.*, pp. 1–19, Jul. 2022, Art. no. 67.
[56] M. Afif, R. Ayachi, Y. Said, and M. Atri, "Deep learning based application for indoor scene recognition," *Neural Process. Lett.*, vol. 51, no. 3, pp. 2827–2837, Mar. 2020.
[57] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 6105–6114.
[58] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4510–4520.
[59] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2818–2826.
[60] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.
[61] S. Jain, U. Singhania, B. Tripathy, E. A. Nasr, M. K. Aboudaif, and A. K. Kamrani, "Deep learning-based transfer learning for classification of skin cancer," *Sensors*, vol. 21, no. 23, p. 8142, Dec. 2021.
[62] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, Mar. 2014.
[63] A. Guernine and M. T. Kimour, "Optimized training for convolutional neural network using enhanced grey wolf optimization algorithm," *Informatica*, vol. 45, no. 5, pp. 731–739, Aug. 2021.
[64] R. Mohakud and R. Dash, "Designing a grey wolf optimization based hyper-parameter optimized convolutional neural network classifier for skin cancer detection," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 34, no. 8, pp. 6280–6291, Sep. 2022.

AHMED MAGDY (Member, IEEE) was born in Cairo, Egypt, in 1986. He received the B.S. degree in electronic and communication engineering and the M.Sc. degree in communication engineering from Helwan University, Egypt, in 2008 and 2013, respectively, and the Ph.D. degree in electrical engineering from the Communication Branch, Minia University, Egypt, in 2017. He is currently an Assistant Professor with the Electrical Engineering Department, Communication Branch, Suez Canal University, Ismailia, Egypt. His research interests include embedded systems, MEMS, image/signal processing AI, the IoT, computer vision, and smart antenna. He is a reviewer of the IET journals.

HADEER HUSSEIN was born in Ismailia, Egypt, in 1989. She received the B.S. degree in computers and control engineering and the M.Sc. degree in computers engineering from Suez Canal University, Ismailia, in 2011 and 2018, respectively. She is currently an Assistant Lecturer with the Electrical Engineering Department, Computers and Control Branch, Suez Canal University. Her research interests include artificial intelligence and computer vision.

REHAB F. ABDEL-KADER received the Ph.D. degree in computer engineering from Auburn University, Auburn, AL, USA, in 2003. She was an Assistant Professor with the Electrical Engineering Department, Georgia Southern University, Statesboro, GA, USA, for two years, before coming back to Egypt. She is currently a Professor and the Vice Dean of Graduate Studies and Research with the Faculty of Engineering, Port Said University. Her current research interests include artificial intelligence and computer vision.

KHALED ABD EL SALAM was born in Egypt, in 1973. He received the B.S., M.Sc., and Ph.D. degrees in computer engineering from Suez Canal University, Ismailia, Egypt, in 2003 and 2010, respectively. From 2000 to 2013, he was a Computer Instructor with The Higher Institute for Hotels and Tourism Management "HIHTM" Lucerne, Hurghada, Egypt. From 2013 to 2018, he was an Associate Professor with the Faculty of Engineering, Sinai University, Egypt. Since 2013, he has been an Associate Professor with the Faculty of Engineering, Suez Canal University.