



# Automated deep learning approach for classification of malignant melanoma and benign skin lesions

Wessam Salma<sup>1</sup> · Ahmed S. Eltrass<sup>2</sup>

Received: 14 December 2020 / Revised: 9 March 2021 / Accepted: 4 April 2022 /

Published online: 14 April 2022

© The Author(s) 2022

## Abstract

Skin cancer becomes a significant health problem worldwide with an increasing incidence over the past decades. Due to the fine-grained differences in the appearance of skin lesions, it is very challenging to develop an automated system for benign-malignant classification through images. This paper proposes a novel automated Computer Aided Diagnosis (CAD) system for skin lesion classification with high classification performance using accuracy low computational complexity. A pre-processing step based on morphological filtering is employed for hair removal and artifacts removal. Skin lesions are segmented automatically using Grab-cut with minimal human interaction in HSV color space. Image processing techniques are investigated for an automatic implementation of the ABCD (asymmetry, border irregularity, color and dermoscopic patterns) rule to separate malignant melanoma from benign lesions. To classify skin lesions into benign or malignant, different pretrained convolutional neural networks (CNNs), including VGG-16, ResNet50, ResNetX, InceptionV3, and MobileNet are examined. The average 5-fold cross validation results show that ResNet50 architecture combined with Support Vector Machine (SVM) achieve the best performance. The results also show the effectiveness of data augmentation in both training and testing with achieving better performance than obtaining new images. The proposed diagnosis framework is applied to real clinical skin lesions, and the experimental results reveal the superior performance of the proposed framework over other recent techniques in terms of area under the ROC curve 99.52%, accuracy 99.87%, sensitivity 98.87%, precision 98.77%, *F1*-score 97.83%, and

---

✉ Ahmed S. Eltrass  
ahmed.eltrass@alexu.edu.eg

Wessam Salma  
Wessam.salama@pua.edu.eg

<sup>1</sup> Basic Science Department, Faculty of Engineering, Pharos University, Alexandria, Egypt

<sup>2</sup> Electrical Engineering Department, Faculty of Engineering, Alexandria University, Alexandria, Egypt

consumed time 3.2 s. This reveals that the proposed framework can be utilized to help medical practitioners in classifying different skin lesions.

## 1 Introduction

Skin cancer has been increasing among men and women worldwide for many decades [32]. There were approximately 76,250 new cases of melanoma and approximately 8,790 new melanoma-related deaths during 2012 in the United States [51]. In Brazil, it is estimated that for the biennium of 2018–2019, there were 165,580 new cases of non-melanoma skin cancer [34]. Skin cancer spreading results from many factors such as long longevity of the population, people being exposed to the sun, and early detection of skin cancer. Dermoscopy, a noninvasive skin imaging technique, is one of the most effective ways for early identification of skin cancer. The appearance of skin lesions in dermoscopic images may change significantly according to the skin condition. In addition, different artifact sources, including hair, skin texture, and air bubbles may lead to misidentifying the boundary between the skin lesions and the surrounding healthy skin.

Despite the effectiveness of dermoscopic diagnosis for skin cancer, it is very difficult for expert dermatologists to provide an accurate classification of malignant melanoma and benign skin lesions for large number of dermoscopic images. Hence, it is very necessary to build up a non-invasive Computer Aided Diagnosis (CAD) system for skin lesion classification. The CAD system generally comprises of four main steps: image pre-processing, segmentation, feature extraction, and classification. Note that each step significantly affects the classification performance of the whole CAD system [50]. Therefore, efficient algorithms should be employed in each step to achieve high diagnosis performance.

Various studies examined distinct machine learning methods for diagnosis of different types of cancer [7, 11, 33, 45]. The majority of these studies used classifiers trained on a set of hand-crafted features captured from the images. Most of machine learning techniques require high computational time for accurate diagnosis and their performance depends on the selected features that characterize the cancerous region. Deep learning techniques and Convolutional Neural Networks (CNNs) have become important for automated diagnosis of different types of cancer [3]. Deep learning has achieved impressive results in image classification applications. In image classification tasks, transfer learning [29] and data augmentation [38] are employed to overcome the lack of data and to reduce the computational and memory requirements.

Transfer learning is an effective tool for classification problems with few available datasets like medical imaging applications. Instead of training a CNN from scratch, which would require a huge amount of data and a high computational cost, it is computationally efficient to employ a pre-trained CNN architecture and just fine-tune its performance to accelerate the operation. Various pre-trained CNNs, including AlexNet, Inception, ResNet, and DenseNet [8, 37] were trained on the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) dataset. Some studies employed distinct pre-trained CNNs and reported reasonable diagnosis performance for skin cancer [2, 13, 26, 30, 31, 36].

Data augmentation is a technique for enhancing the size of the input data by developing new data from the original data [38]. Image augmentation techniques can be employed to defeat the lack of skin cancer datasets. There are many data augmentation strategies, including rotation, scaling, random cropping, and color modifications. Data augmentation is broadly employed with pre-trained CNN architectures. In [30], the performance of skin lesion classification was investigated with 13 data augmentation techniques trained on three CNNs

(Inception-v4, ResNet, and DenseNet), and the results revealed the positive impact of using data augmentation in both training and testing phases. An automatic skin lesion classification system based on the Alex-Net CNN architecture was proposed by [13]. The architecture weights were fine-tuned, and the datasets are augmented by fixed and random rotation angles, yielding average accuracy of 95.91%.

In [26], a deep learning approach based on neural network ensemble was employed to classify skin lesions using dermoscopic images. The melanoma classification system of [31] was developed for extracting asymmetry, border irregularity, color variation, diameter, and texture of the skin lesion and classifying malignant melanoma and benign skin lesions using Support Vector Machine (SVM) classifier. In [2], a deep CNN prediction model based on new regularizer technique was employed for skin lesion classification with an accuracy of 97.49%. In the work of [36], a Gabor wavelet-based CNN model for skin lesion diagnosis was developed by employing Gabor filters on skin images to obtain detailed characteristics of the skin lesions and then modeling these details using deep CNN models. Decision fusion based on the sum rule was used for classifying the skin lesions with an accuracy of 83%.

The majority of exiting skin lesion diagnosis systems report reasonable classification results for discriminating malignant melanoma from benign lesions. However, it is still very difficult to develop an efficient automated framework that can provide very accurate diagnosis results, while running in real-time using low hardware specifications. The aim of this study is to propose an accurate diagnosis system for skin lesion classification in dermoscopic images. The proposed approach is based on employing efficient image processing techniques for pre-processing, data augmentation, and image segmentation along with investigating different pre-trained CNN architectures for skin lesion classification. The proposed framework is examined on the most common dermoscopic image databases, the first one is the International Skin Imaging Collaboration (ISIC) 2017 [15], and the second one is the MNIST: HAM10000 [17].

The main contributions of this work can be summarized as follows: (1) A new real-time automated CAD system is proposed for skin lesion classification with high performance using low computational complexity. (2) The proposed framework is based on new pre-processing techniques for hair removal and lesion identification, Grab-cut in HSV color space for lesion segmentation, and automatic detection of the ABCD features for discriminating malignant melanoma from benign lesions. (3) Different pretrained CNNs, including VGG-16, ResNet50, ResNetX, InceptionV3, and MobileNet are examined for skin lesion classification. (4) The impact of applying data augmentation to all pretrained CNN models under investigation (VGG-16, ResNet50, ResNetX, InceptionV3, and MobileNet) is examined using several evaluation metrics, including area under the ROC curve, accuracy, sensitivity, precision, *F1*-score, and computational time using 5-fold cross-validation. The results revealed the positive influence of utilizing data augmentation in improving the diagnosis performance. Extensive experiments have been performed to evaluate the performance of the proposed framework, and experimental results reveal the superior performance of the proposed approach over other exiting state-of-the-art diagnosis systems in terms of the classification rate, the sensitivity, the specificity, and the computational efficiency. This demonstrates that the proposed framework represents an efficient tool to assist dermatologists in real-time skin lesion classification.

The remaining parts of this paper are structured as follows. Section 2 presents the techniques used in pre-processing, segmentation, feature extraction, and classification.

Section 3 shows the experimental results of the proposed skin lesion classification system using real dermoscopic images, followed by a comparison between the proposed framework and other recent methods. In Section 4, the conclusions are presented.

## 2 Methodology

The proposed framework for skin lesion classification comprises of four main stages: **image pre-processing, segmentation, feature extraction, and classification**. The structure of the proposed approach is shown in Fig. 1.

### 2.1 Image pre-processing

The pre-processing step is used to eliminate all artifact sources such as bubbles and hair from skin images. This step is very essential before any segmentation or feature extraction to provide accurate diagnosis. A new approach based on morphological filtering is investigated to remove the hair from the skin images [35]. This technique is divided into two mainly steps. First, the color images are converted into grayscale versions using the weighted process of [47], where the grayscale image is obtained from the RGB color space by the relation:  $\text{grayscale} = 0.3R + 0.59G + 0.11B$ . Second, the hair contour of the grayscale images is identified using morphological black-hat transformation [52]. Finally, the inpainting function is used to create a mask based on the Fast-Marching Method (FMM) [42]. In the current work, when the pixel value is lower than the threshold, it will be set to 0, otherwise it will be set to 1.

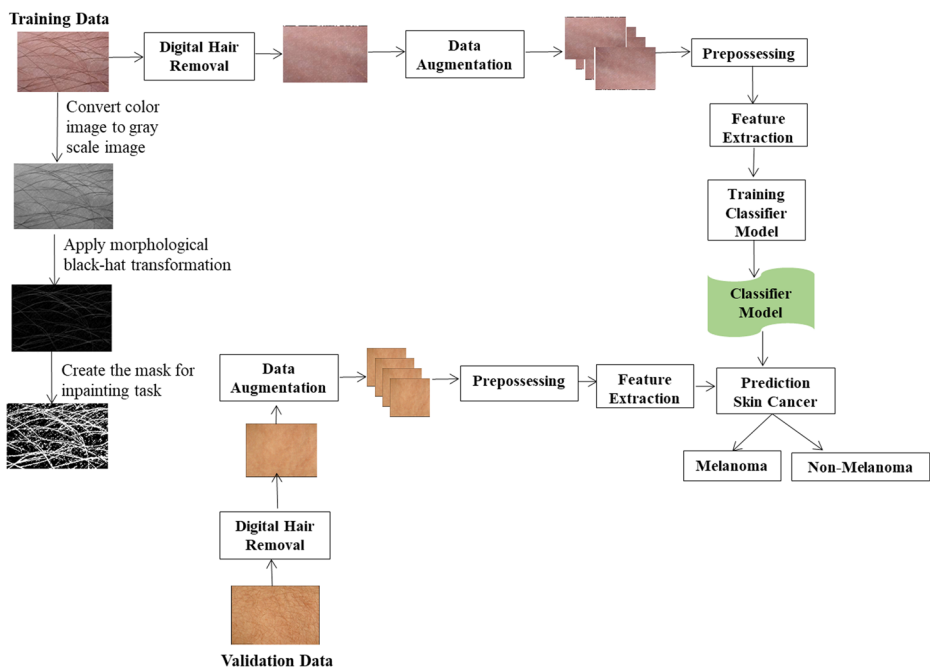


Fig. 1 The structure of the proposed skin cancer CAD system

After the pre-processing step, each pre-processed skin image is augmented into four images by rotating the input image into four directions of  $0^\circ$ ,  $90^\circ$ ,  $180^\circ$ , and  $270^\circ$ . Data augmentation is investigated in the current work to increase the data size, generate new data from the original input data, and overcome the lack of the tagged images.

## 2.2 Grab-cut skin lesion segmentation

After the pre-processing step, the images are segmented automatically using Grab-cut segmentation technique [22]. In this technique, the boundary regions are identified, i.e., region between the lesion, background, and the difference within the lesion. All skin images are converted into HSV color space to distinguish image intensity from color information which is a significant step for the adaptive histogram equalization process as shown in Fig. 2. Using the Gaussian Mixture Model (GMM) [54], the foreground region is extracted from the Grab-cut mask resulted from the adaptive histogram equalization process. Any region beyond this rectangle will be taken as background. The segmentation results are quantitatively evaluated by determining the overlap with the ground truth. Generally, Jaccard Coefficient (JC) is the most frequently used evaluation metric in image segmentation analysis. The JC measures pixels similarity between the ground truth and the segmented image to examine which pixels are matched and which are unmatched. It is defined as the intersection over union between index segmentation results and ground truth [16]. Its value lies between 0 and 1, with 1 showing perfect overlap and 0 indicating no overlap. The segmented images with their JC values are shown in Table 1.

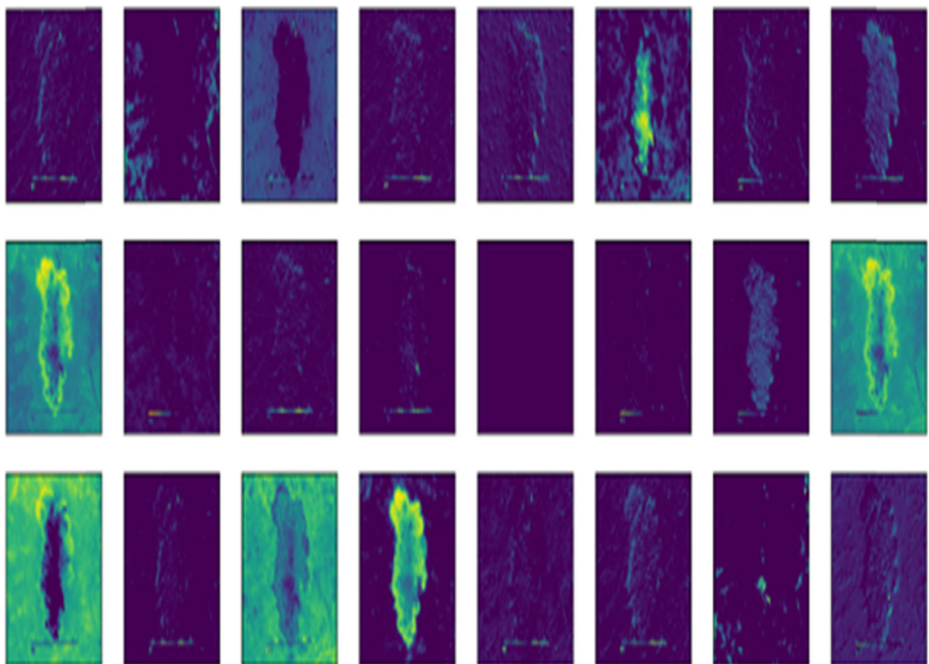
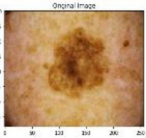

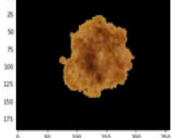
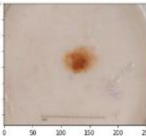
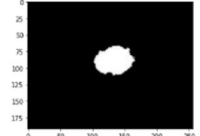
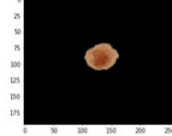


Fig. 2 Conversion from RGB into HSV color space

**Table 1** Segmentation results using Grab-cut technique


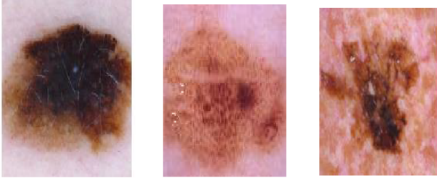




Original image	Ground Truth Mask	Segmented Image	JC
			0.89
			0.82

### 2.3 Feature extraction

To distinguish malignant melanoma from benign skin lesions, the main features of skin lesions should be extracted. The ABCD rule of dermoscopy [28] is investigated in the current study for skin lesion identification. The ABCD features are extracted from the intrinsic physical attributes of skin lesions, including asymmetry, border irregularity, color variation, and dermoscopic patterns. The ABCD rule is used to characterize the geometrical and structural lesion properties and to separate melanoma from benign lesions because of its effectiveness, efficiency, and simplicity of performance and implementation. Table 2 shows the results of ABCD rule of dermoscopy. The four main components of this method can be summarized as follows [19]:

- **Asymmetry A:** Axes of asymmetry are obtained by drawing two orthogonal axes crossing at the center of gravity of the lesion. For a given axis, the determination of whether a lesion is symmetric or not is identified based on the symmetry along that axis for all measures, namely color, brightness, and shape. The asymmetry score is zero if there is no asymmetry in both axes. A score of one is given if there is asymmetry by one axis, and a score of two is given in the case of asymmetry by both axes.
- **Border B:** The lesion is partitioned into eight parts (slices). A border score is assigned according to the existence of sharp or gradual cut-off at the periphery of the skin lesions. A score of zero is given for regular periphery, while a score of one is given for irregular boundary. The border score varies between zero and eight according to the number of slices with irregular periphery. Vector product and inflexion point descriptors are utilized to extract the number of peaks, valleys, and straight lines at lesion edges corresponding to clear irregular and small irregular peripheries, respectively [10].
- **Color C:** It is one of the most important discriminators of malignant melanoma from benign lesions. Lesions with at least one of six colors, namely white, red, black, light brown, dark brown, and blue-gray, are probably melanomas. The existence of one suspicious color is confirmed when the number of pixels concerning that color surpasses by 5% of the total number of pixels of the lesion [43]. In a lesion, more than one suspicious

**Table 2** Feature extraction using ABCD rule of dermoscopy

<b>Asymmetry</b>			
<b>Border structure</b>			
<b>Color variation</b>	 109	 114	 118
<b>Dermoscopic structures</b>			

color may occur, and the color score is increased by one for each existing color. The color score ranges from one to six.

- Dermoscopic structures D: The existence of five distinct structures, namely network, structureless areas, branched streaks, dots, and globules, within the lesion is investigated. The existence of each structure within the lesion increments the score by one. The minimum structure score is zero and the maximum structure score is five.

## 2.4 Classification with CNN

In this step, the resulted ABCD dermoscopy images are fed into a pretrained deep CNN architecture. Different pretrained CNN architectures, including VGG-16 [39], ResNet50 [18], ResNetX [9], InceptionV3 [41], and MobileNet [14] are investigated for skin lesion classification. Note that designing a new deep CNN network from scratch is hard because (i) it needs a huge amount of labeled training data and high experience to guarantee efficient convergence; (ii) the process is time-consuming and tedious; (iii) it may suffer from overfitting problems. Instead of training a CNN from scratch, it is computational efficient to employ a pre-trained CNN model and just fine-tune its performance to accelerate the operation. All pre-trained CNNs considered in the current study were trained on a huge dataset, namely ILSVRC. They were trained on approximately 1.2 million training images with another 50,000 images for validation and 100,000 images for testing. In this work, we use the weights pretrained on



ImageNet to fine-tune the model. This makes the pretrained network can learn new task faster and easier than the network which is trained from scratch using randomly initialized weight.

For all pretrained CNN architectures under investigation, classification results are examined, and the results reveal that VGG-16 and ResNet50 architectures achieve the best performance. In this work, a batch size of 40/16, learning rate of  $10^{-3}/10^{-4}$ , and momentum of 0.8/0.9 are, respectively, the best values for yielding the highest diagnosis performance of ResNet50 and VGG-16 architectures. The detailed layers information of ResNet50 and VGG-16 architectures are shown in Tables 3 and 4, respectively. The proposed framework is enhanced by combining the linear kernel SVM classifier [4] with both VGG-16 and ResNet50 architectures and investigating their classification results.

To authenticate the diagnosis performance and avoid any bias, the k-fold cross-validation technique is employed [40]. Usually, the value of k is aprioristically selected and fixed. Generally, increasing the value of k from 2 to 10 leads to bias decrease, variance increase in the test error rate estimation, and more computational burden [25]. Note that the computational time tends to follow an increasing exponential distribution with the increase of k. Larger values of k are usually employed to use a larger number of samples for training at the expense of error estimation degradation of the classifier [12]. In the current study, the training data is sufficient to use 5-fold cross validation which allows reducing the variance estimate of error prediction and minimizing the computational cost. This is very critical for practical implementation of the proposed CAD system in real-time applications. Also, obtaining a tight and rigorous estimation of the classifier error allows validating, in a statistical sense, the reliability of the model [12]. Moreover, when employing the 5-fold cross validation, larger fraction of samples can be kept for the classification phase. For the data under investigation, the 5-fold cross validation represents a good choice to compromise between the percentage of data used for training, and the rigor of the estimated error.

**Table 3** Summary of the CNN layers for the ResNet50 structure

Layer No.	Type	Size	Kernel/Stride	Activation
	Input	224*224*3	-----	-----
1	Convolution	64*3*3	2	ReLU
2	Convolution	1*1*64	2	ReLU
	Convolution	3*3*64		
	Convolution	1*1*256		
3	Convolution	1*1*128	2	ReLU
	Convolution	3*3*128		
	Convolution	1*1*512		
4	Convolution	1*1*256	2	ReLU
	Convolution	3*3*256		
	Convolution	1*1*1024		
5	Convolution	1*1*512	2	ReLU
	Convolution	3*3*512		
	Convolution	1*1*2048		
	Output	2	----	SVM



**Table 4** Summary of the CNN layers for the VGG-16 structure

Layer No.	Type	Size	Kernel size	Activation
1	Input	224*224*3	----	----
	Convolution	224*224*64	-----	ReLU
	Convolution	224*224*64	-----	
	MaxPooling	112*112*64	2*2	
2	Convolution	112*112*128	-----	ReLU
	Convolution	112*112*128	-----	
	MaxPooling	56*56*128	2*2	
	Convolution	56*56*256	-----	
3	Convolution	56*56*256	----	ReLU
	Convolution	56*56*256	----	
	Convolution	56*56*256	----	
	MaxPooling	28*28*256	2*2	
4	Convolution	28*28*512	----	ReLU
	Convolution	28*28*512	----	
	Convolution	28*28*512	----	
	MaxPooling	14*14*512	2*2	
5	Convolution	14*14*512	----	ReLU
	Convolution	14*14*512	----	
	Convolution	14*14*512	----	
	MaxPooling	14*14*512	2*2	
6	FC	4096*1	-----	ReLU + Dropout (0.5)
7	FC	4096*1	-----	ReLU + Dropout(0.5)
8	FC	4096*1	-----	ReLU + Dropout(0.5)
	Output	2		SVM

### 3 Results

To examine the usefulness of the proposed skin lesion classification approach, 580 images (310 benign and 270 malignant) from the ISIC 2017 database and 350 images (215 benign and 135 malignant) from the MNIST: HAM10000 database are investigated. The data under investigation covers distinct cases of abnormality, patient age, and patient gender as shown in Table 5. In order to increase the total number of images and allow more data for training and testing, each image is augmented into four images of different directions of rotation, providing a total number of 2320 image from the ISIC 2017 database and 1400 image from the MNIST: HAM10000

**Table 5** Distribution of images taken from the ISIC 2017 and MNIST: HAM10000 databases

ISIC2017 Database: 580 Images (270 Malignant and 310 Benign)				
Patients Age (Years)				
30–40	41–50	51–60	61–70	> 70
90	120	210	100	60
Patients Gender (Male and Female)				
Male	236			
Female	344			
HAM10000 Database: 350 Images (135 Malignant and 215 Benign cases)				
Patients Age (Years)				
30–40	41–50	51–60	61–70	> 70
60	105	65	80	40
Patients Gender (Male and Female)				
Male	193			
Female	157			

database. Evaluation of image classification depends mainly on computing four parameters: the number of true positives ( $TP$ ), true negatives ( $TN$ ), false negatives ( $FN$ ), and false positives ( $FP$ ). The classification performance is identified in terms of Accuracy ( $ACC$ ), Sensitivity ( $Se$ ) or recall, Precision ( $Pr$ ),  $F1$ -score, area under the ROC curve ( $AUC$ ), and computational time.  $ACC$  is a metric of true predictions,  $Pr$  is the fraction of detected malignant cases that agree with the ground truth, and  $Se$  is the fraction of true malignant cases that are identified malignant.  $F1$ -score is the harmonic mean of  $Pr$  and  $Se$ .  $AUC$  measures the entire two-dimensional area underneath the entire ROC curve and it gives a total performance evaluation across all possible classification thresholds. All these metrics are defined as follows:

$$ACC = (TP + TN) / (TP + FP + TN + FN) \quad (1)$$

$$Pr = TP / (TP + FP) \quad (2)$$

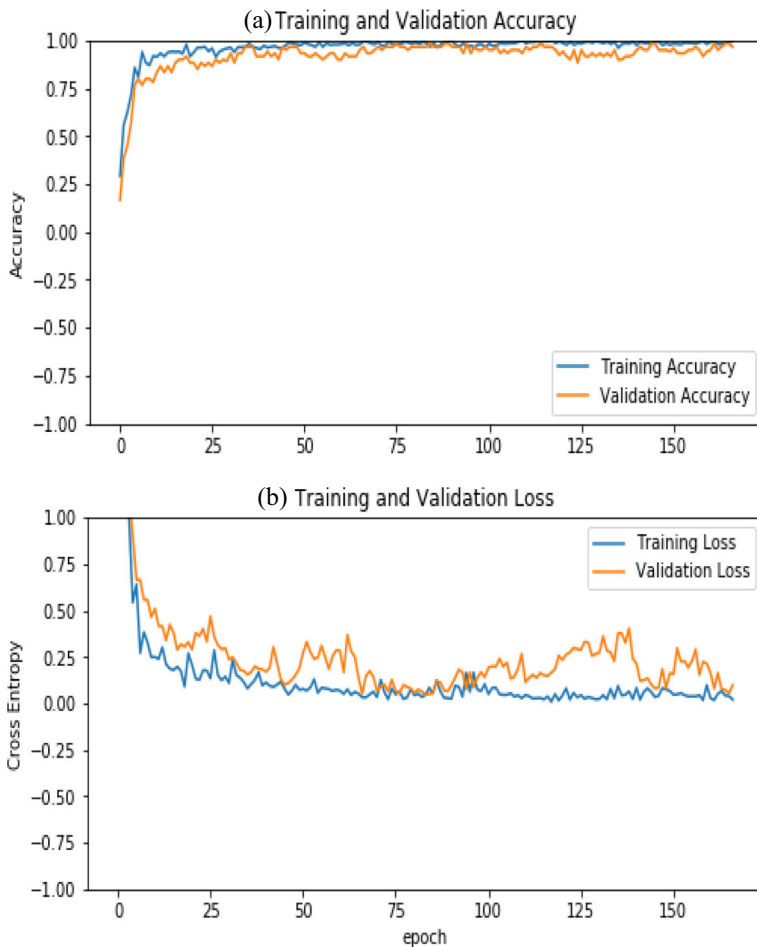
$$Se = TP / (TP + FN) \quad (3)$$

$$F1 = 2(Pr \times Se) / (Pr + Se) \quad (4)$$

All images under investigation (2320 augmented images from the ISIC 2017 and 1400 augmented images from the MNIST: HAM10000) are divided into 2 main partitions: 1120 for training and validation and 280 for testing using the MNIST: HAM10000 database, and 1856 for training and validation, and 464 for testing using the ISIC 2017 database. To authenticate the diagnosis performance and avoid any possible bias, the 5-fold cross-validation is utilized, where the training and validation data (1856 from MNIST: HAM10000 database and 1856 from the ISIC 2017 database) are partitioned into five equal sized groups, each time one group is retained as a validation set and the remaining four groups are retained as a training set. The classification metrics are computed for each fold relevant to the test dataset for both datasets.

For all pretrained CNN architectures under investigation (VGG-16, ResNet50, ResNetX, InceptionV3, and MobileNet), training and validation loss curves along with training and validation accuracy curves are investigated, and the results reveal that VGG-16 and ResNet50 architectures achieve the best performance. For example, the training progress versus the number of epochs for the ResNet50 architecture is shown in Fig. 3a. The ResNet50 network has faster and better training advance than other CNN networks. Also, it achieves higher performance than other CNN architectures in terms of the validation accuracy as shown in Fig. 3a. The VGG-16 architecture achieves comparable performance with ResNet50 network. Therefore, a critical comparison is made between VGG-16 and ResNet50 architectures for skin lesion classification under various parameter regimes.

Table 6 shows the 5-fold benign-malignant classification results and standard deviations of the proposed system using VGG-16 and ResNet50 architectures for the ISIC 2017 database with data augmentation. The 5-fold average classification results of  $AUC$ ,  $ACC$ ,  $Pr$ ,  $Se$ , and  $F1$ -score are calculated by averaging the 5-fold classification results (see Table 6). The average



**Fig. 3** Training and validation (a) accuracy and (b) loss versus number of epochs for the ResNet50 architecture

classification results for each metric were calculated for all CNN architectures under investigation (VGG-16, ResNet50, ResNetX, InceptionV3, and MobileNet), showing the better performance of ResNet50 architecture in terms of all metrics (*ACC*, *AUC*, *Pr*, *Se*, and *F1*-score). Similar results are obtained for all CNN architectures using MNIST: HAM10000 database, and the results show the higher performance of ResNet50 architecture as shown in Table 7.

We further enhance the proposed system by combining the kernel SVM classifier with all CNN architectures and investigating their classification results. Table 7 shows the average 5-fold benign-malignant classification results and standard deviations of VGG-16, ResNet50, ResNetX, InceptionV3, and MobileNet architectures with and without SVM for both ISIC 2017 and MNIST: HAM10000 databases with and without data augmentation. It can be noted that the results of either VGG-16 or ResNet50 architecture with SVM are better than the standalone VGG-16 or ResNet50. Also, the performance of ResNet50 architecture combined with SVM is superior to VGG-16 architecture with SVM in terms of average *ACC* (99.87% versus 98.66%), *AUC* (99.52% versus 97.63%), *Se* (98.87% versus 95.49%), *Pr* (98.77%

**Table 6** Average 5-fold classification results and standard deviation for ResNet50 and VGG-16 using 2320 images from the ISIC 2017 database

ResNet50 Architecture					
Fold	ACC	AUC	Se	Pr	F1-score
1	95.62%	93.42%	95.98%	93.94%	92.42%
2	95.78%	93.11%	94.88%	95.58%	94.98%
3	95.84%	92.44%	94.98%	95.68%	93.98%
4	95.75%	93.45%	93.99%	93.87%	92.99%
5	96.56%	92.47%	93.89%	94.99%	93.87%
Average 5-Fold Results	95.91% ± 0.26	92.97% ± 0.29	94.74% ± 0.31	94.81% ± 0.28	93.64% ± 0.27
VGG-16 Architecture					
Fold	ACC	AUC	Se	Pr	F1-score
1	97.75%	93.59%	94.84%	95.54%	94.98
2	96.45%	92.87%	92.98%	93.45%	93.95%
3	96.61%	92.89%	93.45%	91.81%	92.98%
4	96.77%	91.98%	94.74%	92.97%	91.97%
5	95.93%	92.87%	92.85%	93.89%	93.85%
Average 5-Fold Results	96.70% ± 0.36	92.84% ± 0.34	93.77% ± 0.31	93.53% ± 0.35	93.54% ± 0.30

**Table 7** Average 5-fold classification results and standard deviations for VGG-16, ResNet50, ResNetX, InceptionV3, and MobileNet architectures with and without SVM for both ISIC 2017 and MNIST: HAM10000 databases with and without data augmentation

ISIC 2017 database with data augmentation					
Model	ACC	AUC	Se	Pr	F1-score
ResNet50+SVM	99.87% ± 0.21	99.52% ± 0.27	98.87% ± 0.25	98.77% ± 0.23	97.83% ± 0.21
ResNet50	95.91% ± 0.26	92.97% ± 0.29	94.74% ± 0.31	94.81% ± 0.28	93.64% ± 0.27
VGG16+SVM	98.66% ± 0.27	97.63% ± 0.29	95.49% ± 0.31	96.79% ± 0.26	95.77% ± 0.25
VGG16	96.70% ± 0.36	92.84% ± 0.34	93.77% ± 0.31	93.53% ± 0.35	93.54% ± 0.30
ResNetX+SVM	93.65% ± 0.21	93.82% ± 0.26	93.52% ± 0.24	93.64% ± 0.22	93.69% ± 0.23
ResNetX	92.68% ± 0.14	92.69% ± 0.17	92.78% ± 0.15	92.98% ± 0.11	92.79% ± 0.19
InceptionV3+SVM	90.89% ± 0.18	90.78% ± 0.17	91.21% ± 0.13	90.89% ± 0.18	90.99% ± 0.16
InceptionV3	87.58% ± 0.32	87.34% ± 0.35	87.34% ± 0.31	88.02% ± 0.36	87.36% ± 0.36
MobileNet+SVM	85.68% ± 0.26	85.83% ± 0.29	86.02% ± 0.23	85.89% ± 0.28	86.57% ± 0.29
MobileNet	83.99% ± 0.38	84.98% ± 0.32	85.03% ± 0.31	84.68% ± 0.33	84.98% ± 0.35
ISIC 2017 database without data augmentation					
ResNet50+SVM	96.98% ± 0.20	96.55% ± 0.21	96.96% ± 0.23	96.87% ± 0.21	95.98% ± 0.20
VGG16+SVM	94.87% ± 0.28	93.96% ± 0.29	92.48% ± 0.30	92.93% ± 0.27	93.83% ± 0.28
ResNetX+SVM	91.48% ± 0.14	91.65% ± 0.18	91.69% ± 0.18	91.88% ± 0.15	91.98% ± 0.13
InceptionV3+SVM	89.78% ± 0.11	89.65% ± 0.18	89.79% ± 0.16	88.99% ± 0.15	89.32% ± 0.19
MobileNet+SVM	84.32% ± 0.21	83.99% ± 0.22	84.65% ± 0.24	84.68% ± 0.26	84.83% ± 0.27
MNIST: HAM10000 database with data augmentation					
ResNet50+SVM	98.79% ± 0.35	97.91% ± 0.33	97.78% ± 0.31	96.86% ± 0.29	97.45% ± 0.27
ResNet50	96.78% ± 0.37	96.43% ± 0.36	95.43% ± 0.33	93.86% ± 0.31	93.97% ± 0.29
VGG16+SVM	94.74% ± 0.39	94.88% ± 0.35	96.14% ± 0.37	93.98% ± 0.32	92.45% ± 0.31
VGG16	92.76% ± 0.35	92.82% ± 0.31	92.69% ± 0.33	91.86% ± 0.29	92.97% ± 0.26
ResNetX+SVM	92.62% ± 0.15	92.32% ± 0.14	92.69% ± 0.11	92.65% ± 0.15	92.98% ± 0.15
ResNetX	93.21% ± 0.37	92.99% ± 0.36	93.21% ± 0.31	93.26% ± 0.38	93.12% ± 0.35
InceptionV3+SVM	88.88% ± 0.15	87.99% ± 0.13	88.76% ± 0.11	88.68% ± 0.15	88.87% ± 0.17
InceptionV3	86.88% ± 0.37	86.99% ± 0.36	86.74% ± 0.33	86.69% ± 0.31	87.01% ± 0.35
MobileNet+SVM	84.84% ± 0.23	84.78% ± 0.21	84.89% ± 0.29	84.99% ± 0.23	84.87% ± 0.28
MobileNet	83.56% ± 0.15	83.69% ± 0.11	83.89% ± 0.13	83.63% ± 0.19	83.78% ± 0.12
MNIST: HAM10000 database without data augmentation					
ResNet50+SVM	95.97% ± 0.32	94.98% ± 0.30	94.86% ± 0.30	95.83% ± 0.28	95.97% ± 0.27
VGG16+SVM	94.89% ± 0.35	94.96% ± 0.33	93.87% ± 0.32	92.78% ± 0.31	92.68% ± 0.28
ResNetX+SVM	89.89% ± 0.14	90.01% ± 0.19	90.28% ± 0.18	89.99% ± 0.15	90.07% ± 0.13
InceptionV3+SVM	86.87% ± 0.12	86.79% ± 0.19	86.96% ± 0.18	86.76% ± 0.15	86.86% ± 0.13
MobileNet+SVM	82.93% ± 0.23	82.77% ± 0.20	82.84% ± 0.26	82.65% ± 0.21	82.98% ± 0.24

versus 96.79%), and *F1*-score (97.83% versus 95.77%) for ISIC 2017 database. The same superior performance of ResNet50 architecture with SVM is achieved for the MNIST: HAM10000 database. Note that the ResNet50 architecture with SVM provides smaller standard deviation values than VGG-16 with or without SVM, showing its reliability and consistency in skin lesion classification. In comparing the classification results with and without data augmentation for both databases, results reveal that data augmentation allows increasing the size of data used for training and testing, and hence enhancing the classification results for all cases.

### 3.1 Performance analysis

The proposed ResNet50 CNN architecture combined with SVM for skin lesion classification is compared with other recent systems [1, 2, 5, 6, 13, 19, 20, 23, 24, 30, 31, 36, 46, 48, 49, 53], and the results are shown in Table 8. The average classification results using 5-fold cross-validation are presented to validate the performance of the proposed system when comparing it with other techniques. Table 8 reveals that we investigated larger number of lesion images than most of other techniques under comparison. It can be noted that the proposed approach outperforms other diagnosis systems in terms of *ACC* (99.87% for ISIC 2017 and 98.97% for HAM10000) and *AUC* (99.52% for ISIC 2017 and 97.91% for HAM10000). The proposed system achieves the highest *Se* in comparing with other techniques [1, 2, 13, 19, 20, 24, 31, 46, 48, 49, 53], revealing its effectiveness in classifying patients with skin cancer. Note that sensitivity is a more important evaluation metric than other metrics for dermatologists and clinicians because low sensitivity or high rate of false negative (malignant identified as benign) may result in death. In this study, very few malignant lesions are misclassified, and consequently the proposed approach has superior sensitivity results. Table 8 demonstrates the superior diagnosis results of the proposed CAD system over recent state-of-the-art methods.

Different pretrained CNN architectures, including VGG-16, ResNet50, ResNetX, InceptionV3, and MobileNet were investigated for skin cancer diagnosis, and the results show that ResNet50 architecture achieves the best performance. A detailed comparison between all these CNN architectures is made, and the results demonstrate the better performance of ResNet50 architecture with low computational requirements. It is well-known that there are various pretrained CNN structures [27], and we have already examined most of them. More deeper CNN models mean more computational burden. This is not practical for real-time diagnosis systems. Also, the use of more deep layers increases the number of free parameters, which may lead to over-fitting issues and performance declination. In the current study, the data variability is smaller than other image classification applications, and hence deeper CNNs are not computationally effective for real-time diagnosis. The CNN models selected for this study represent proper trade-off between speed and accuracy.

The overall average computational time of the proposed ResNet50 architecture combined with SVM for skin lesion classification is around 3.23 s on a Kaggle Notebook GPU cloud (2 CPU cores and 13 GB RAM) using Python software. In comparing with other studies under comparison, most of them did not report any computational time analysis. In [53], they reported the computational time using NVIDIA GTX Titan XP GPU which was used for training the model in round 30 h with an average test time of 0.02 s per patch. Despite the low computational time of [53] in comparing to the proposed approach, it requires high hardware requirements and provides low classification performance (see Table 8). The proposed approach achieves high diagnosis performance, while achieving reasonable computational

**Table 8** Comparison between the proposed diagnosis system and several recent classification methods based on different image processing techniques and CNN architectures using different datasets

Reference	Number of lesion images	Databases	AUC %	ACC %	Se %
The proposed diagnosis system	2320	ISIC 2017	99.52	99.87	98.87
	1400	HAM10000	97.91	98.79	97.78
Pretrained CNNs (Inception-v4, ResNet, and DenseNet) with different data augmentation techniques [30], 2018.	2000	ISIC 2017	88.20	-----	-----
Pretrained Alex-Net architecture with data augmentation [13], 2019.	2000	ISIC 2017	-----	95.91	88.47
ABCD features, feature selection with principal component analysis (PCA), SVM classifier [31], 2014.	282	Private	88.10	82.2	77.02
Deep CNN prediction model based on new regularizer technique [2], 2019.	1760	ISIC archive	98.00	97.49	94.30
Gabor wavelet-based deep CNN [36], 2019.	2000	ISIC 2017	96.00	83.00	-----
Automatic ABCD scoring of skin lesions, [19], 2016.	200	Private	-----	94.00	91.20
Fully convolutional residual network [49], 2016.	1250	ISBI 2016	-----	94.40	91.10
Optimal color and texture features, majority-voting multi-class SVM [1], 2016.	350	Atlas of Dermoscopy CD	94.00	93.00	94.00
Pre-processing with cytological properties, CNN with 5-fold cross validation [48], 2016.	1760	ISIC archive	84.70	86.60	80.90
Deep CNN with data augmentation [6], 2016.	211	ISIC 2017	-----	88.00	-----
CNN model based on attention residual learning [53], 2019.	2750	ISIC 2017	87.50	85.00	65.80
Ant colony-based segmentation, geometric and texture features, ANN classifier [5], 2017.	172	DermIS and DermQuest	-----	93.60	-----
Probabilistic distribution segmentation, entropy-based method feature selection, multi-class SVM [20], 2018.	200 290 2750	PH2, ISIC archive ISBI 2016 & 2017	----- ----- 96.00	97.50 97.75 93.2	96.67 96.60 93.00
Deep CNN with data augmentation [46], 2017.	900	ISBI 2016	-----	83.60	76.00
Deep CNN combined with hand-crafted features and local binary patterns [24], 2016.	1279	ISIC archive	78.00	82.60	53.33
Optimized deep features from well-established CNNs and from different abstraction levels [23], 2019.	2037	ISIC 2017	83.83	-----	-----

efficiency using low hardware specifications. In comparing with the deep learning model of [21] which achieved high ACC of 94.5%, the computational efficiency of the proposed framework outperforms it in terms of testing time (3.23 s versus 6.9 s). This elucidates the usefulness of the proposed framework in improving the skin lesion classification performance through real-time systems.

Despite the high performance of the proposed skin lesion classification system, there are few limitations that can be investigated in the future. First, the number of pre-trained CNN architectures under investigation can be increased to incorporate more recent and advanced pre-trained models. Second, the process of extracting comprehensive features such as ABCD features may include some practical limitations like misclassifying melanoma of homogeneous color and regular shape. Future work on this topic may involve the use of new deep learning techniques for further improvement of artifacts removal and structures detection. Also, the hardware design and implementation of the proposed CAD system using high dimensional

chaotic circuits [44] can be a goal for future investigation. These points are challenges that will be explored and reported in the near future.

## 4 Conclusions

In this paper, a new CAD system is proposed for automated skin lesion classification using low hardware and software requirements. The proposed approach investigates new pre-processing techniques for hair removal and lesion identification, Grab-cut in HSV color space for lesion segmentation, the ABCD rule for discriminating malignant melanoma from benign lesions, image data augmentation, and pretrained CNN architectures for classification. Different pretrained CNN structures such as VGG-16, ResNet50, ResNetX, InceptionV3, and MobileNet are investigated for skin lesion classification. To investigate the usefulness of the proposed approach, real skin images covering several cases from the ISIC 2017 and HAM10000 databases are examined. The use of data augmentation is investigated for overcoming the problem of data limitation and examining its impact on the performance of benign-malignant classification for all CNN architectures under investigation. The classification performance is identified using several metrics, including *AUC*, *ACC*, *Se*, *Pr*, *F1-score*, and consumed time. Experimental results reveal that the ResNet50 CNN architecture combined with SVM and data augmentation provides very accurate skin lesion classification with an average *AUC* of 99.52%, *ACC* of 99.87%, *Se* of 98.87%, *Pr* of 98.77%, and *F1-score* of 7.83% for the ISIC 2017 datasets, and *AUC* of 97.91%, *ACC* of 98.79%, *Se* of 97.78%, *Pr* of 96.86%, and *F1-score* of 97.45% for the HAM10000 datasets, while requiring low computational time. The results highlighted the positive impact of utilizing data augmentation in enhancing the diagnosis performance. The proposed method is compared with recent skin lesion classification systems, and the results elucidate the superior classification results of the proposed method over other systems under comparison. This demonstrates that the proposed framework can be used to assist less experienced dermatologists and clinicians in classifying various skin lesions.

**Funding** Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Abbas Q, Sadaf M, Akram A (2016) Prediction of dermoscopy patterns for recognition of both melanocytic and non-melanocytic skin lesions. *Computers* 5(3):13. <https://doi.org/10.3390/computers5030013>
2. Albahar MA (2019) Skin lesion classification using convolutional neural network with novel regularizer. *IEEE Access* 7:38306–38313. <https://doi.org/10.1109/ACCESS.2019.2906241>



3. Brinker TJ, Hekler A, Utikal JS, Grabe N, Schadendorf D, Klode J, Berking C, Steeb T, Enk AH (2018) and C. Von Kalle: Skin cancer classification using convolutional neural networks: systematic review. *J Med Internet Res* 20(10):e11936. <https://doi.org/10.2196/11936>
4. Chen B, Lu Y, Pan W, Xiong J, Yang Z, Yan W, Liu L, Qu J (2019) Support vector machine classification of nonmelanoma skin lesions based on fluorescence lifetime imaging microscopy. *Anal Chem* 91(16): 10640–10647. <https://doi.org/10.1021/acs.analchem.9b01866>
5. Dalila F, Zohra A, Reda K, Hocine C (2017) Segmentation and classification of melanoma and benign skin lesions. *Optik* 140:749–761. <https://doi.org/10.1016/j.jilleo.2017.04.084>
6. Demyanov S, Chakravorty R, Abedini M, Halpern A, Gamavi R (2016) Classification of dermoscopy patterns using deep convolutional neural networks. In: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). IEEE, Prague, 364–368. <https://doi.org/10.1109/ISBI.2016.7493284>
7. Eltrass AS, Salama M (2020) Fully automated scheme for computer-aided detection and breast cancer diagnosis using digitised mammograms. *IET Image Process* 14(3):495–505. <https://doi.org/10.1049/iet-ipr.2018.5953>
8. Feng X, Yao H, Zhang S (2019) An efficient way to refine DenseNet. *Signal Image Video Process* 13(5): 959–965. <https://doi.org/10.1007/s11760-019-01433-4>
9. Feng W, Zhang X, Zhao G (2019) ResNetX: a more disordered and deeper network architecture. arXiv preprint arXiv:1912.12165. Available from: <https://arxiv.org/abs/1912.12165>. Accessed 20 Jul 2020
10. Gamavi R, Aldeen M, Bailey J (2012) Computer-aided diagnosis of melanoma using border-and wavelet-based texture analysis. *IEEE Trans Inf Technol Biomed* 16(6):1239–1252. <https://doi.org/10.1109/TITB.2012.2212282>
11. Hardie R, Ali R, Silva D, Kebede TM (2018) Skin lesion segmentation and classification for ISIC 2018 using traditional classifiers with hand-crafted features. arXiv preprint arXiv:1807.07001. Available from: <https://arxiv.org/abs/1807.07001>. Accessed 24 Jul 2020
12. Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning: data mining, inference and prediction, 2nd edn. Springer Science & Business Media, Springer-Verlag New York. <https://doi.org/10.1007/978-0-387-84858-7>
13. Hosny KM, Kassem MA, Foad MM (2019) Classification of skin lesions using transfer learning and augmentation with Alex-net. *PLoS One* 14(5):e0217293. <https://doi.org/10.1371/journal.pone.0217293>
14. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861, 2017. Available from: <https://arxiv.org/abs/1704.04861>. Accessed 10 Aug 2020
15. ISIS Archive [electronic resource] (2020) Kitware, Available: Inc. <https://isic-archive.com/>. Accessed 16 Jan 2020
16. Jaccard P (1912) The distribution of the flora in the alpine zone 1. *New Phytol* 11(2):37–50. <https://doi.org/10.1111/j.1469-8137.1912.tb05611.x>
17. K. Inc (2019) Skin Cancer MNIST: HAM10000. Available: <https://www.kaggle.com/kmader/skin-cancer-mnist-ham10000/version/2>. Accessed 24 Jan 2020
18. Kaiming H, Xiangyu Z, Shaoqing R, Jian S (2016) Deep residual learning for image recognition. *IEEE Conf on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA. <https://doi.org/10.1109/CVPR.2016.90>
19. Kasmi R, Mokrani K (2016) Classification of malignant and benign skin lesions: implementation of automatic ABCD rule. *IET Image Proc* 10(6):448–455. <https://doi.org/10.1049/iet-ipr.2015.0385>
20. Khan MA, Akram T, Sharif M, Shahzad A, Aurangzeb K, Alhusssein M, Haider SI, Altamrah A (2018) An implementation of normal distribution based segmentation and entropy controlled features selection for skin lesion detection and classification. *BMC Cancer* 18(1):638. <https://doi.org/10.1186/s12885-018-4465-8>
21. Khan MA, Sharif M, Akram T, Bukhari SAC, Nayak RS (2020) Developed Newton-Raphson based deep features selection framework for skin lesion recognition. *Pattern Recogn Lett* 129:293–303. <https://doi.org/10.1016/j.patrec.2019.11.034>
22. Li Y, Zhang J, Gao P, Jiang L, Chen M (2018) Grab cut image segmentation based on image region. In: 2018 IEEE 3rd International Conference on Image. Vision and Computing (ICIVC). IEEE, Chongqing, 311–315. <https://doi.org/10.1109/ICIVC.2018.8492818>
23. Mahbod A, Schaefer G, Wang C, Ecker R, Elling I (2019) Skin lesion classification using hybrid deep neural networks. In: ICASSP 2019-IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, Brighton, 1229–1233. <https://doi.org/10.1109/ICASSP.2019.8683352>
24. Majtner T, Yildirim-Yayilgan S, Hardeberg JY (2016) Combining deep learning and hand-crafted features for skin lesion classification. In: 2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA). IEEE, Oulu, 1–6. <https://doi.org/10.1109/IPTA.2016.7821017>
25. Marcot BG, Hanea AM (2020) What is an optimal value of k in k-fold cross-validation in discrete Bayesian network analysis? *Comput Stat* :1–23. <https://doi.org/10.1007/s00180-020-00999-9>

26. Matsunaga K, Hamada A, Minagawa A, Koga H (2017) Image classification of melanoma, nevus and seborrheic keratosis by deep neural network ensemble. arXiv preprint arXiv:1703.03108. <https://arxiv.org/abs/1703.03108>. Accessed 12 Aug 2020
27. Menegola A, Fornaciali M, Pires R, Bittencourt FV, Avila S, Valle E (2017) Knowledge transfer for melanoma screening with deep learning. In 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), pp 297–300, Melbourne, VIC, Australia. <https://doi.org/10.1109/ISBI.2017.7950523>
28. Monisha M, Suresh A, Babu BR, Rashmi MR (2019) Classification of malignant melanoma and benign skin lesion by using back propagation neural network and ABCD rule. *Clust Comput* 22(5):12897–12907. <https://doi.org/10.1007/s10586-018-1798-7>
29. Pan SJ, Yang Q (2009) A survey on transfer learning. *IEEE Trans Knowl Data Eng* 22(10):1345–1359. <https://doi.org/10.1109/TKDE.2009.191>
30. Perez F, Vasconcelos C, Avila S, Valle E (2018) Data augmentation for skin lesion analysis. OR 2.0 context-aware operating theaters, computer assisted robotic endoscopy, clinical image-based procedures, and skin image analysis. Springer, Cham, pp 303–311. [https://doi.org/10.1007/978-3-030-01201-4\\_33](https://doi.org/10.1007/978-3-030-01201-4_33)
31. Ramezani M, Karimian A, Moallem P (2014) Automatic detection of malignant melanoma using macroscopic images. *J Med Signals Sens* 4(4):281. <https://doi.org/10.4103/2228-7477.144052>
32. Rembielak A, Ajithkumar T (2019) Non-melanoma skin cancer—an underestimated global health threat. *Clin Oncol* 31(11):735–737. <https://doi.org/10.1016/j.clon.2019.08.013>
33. Salama MS, Eltrass AS, Elkamchouchi HM (2018) An improved approach for computer-aided diagnosis of breast cancer in digital mammography. 13th Annual IEEE International Symposium on Medical Measurements and Applications, Rome, Italy, 1–5. <https://doi.org/10.1109/McMeA.2018.8438650>
34. Santos MO (2018) Estimate: cancer incidence in Brazil. *Rev Bras Cancerol* 64(1):119–120
35. Serra J (1994) Morphological filtering: an overview. *Sig Process* 38(1):3–11. [https://doi.org/10.1016/0165-1684\(94\)90052-3](https://doi.org/10.1016/0165-1684(94)90052-3)
36. Serte S, Demirel H (2019) Gabor wavelet-based deep learning for skin lesion classification. *Comput Biol Med* 113:103423. <https://doi.org/10.1016/j.combiomed.2019.103423>
37. Shin HC, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers RM (2016) Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans Med Imaging* 35(5):1285–1298. <https://doi.org/10.1109/TMI.2016.2528162>
38. Shorten C, Khoshgoftaar TM (2019) A survey on image data augmentation for deep learning. *J Big Data* 6(1), 60:1–48. <https://doi.org/10.1186/s40537-019-0197-0>
39. Simonyan K, Zisserman A (2017) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. Available from: <https://arxiv.org/abs/1409.1556>. Accessed 1 Aug 2020
40. Stone M (1977) An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion. *J R Stat Soc Ser B Methodol* 39(1):44–47. <https://doi.org/10.1111/j.2517-6161.1977.tb01603.x>
41. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016) Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2818–2826, Las Vegas, NV, USA. <https://doi.org/10.1109/CVPR.2016.308>
42. Telea A (2004) An image inpainting technique based on the fast marching method. *J Graph Tools* 9(1):23–34. <https://doi.org/10.1080/10867651.2004.10487596>
43. Thanh DN, Prasath VS, Hien NN (2020) Melanoma skin cancer detection method based on adaptive principal curvature, colour normalisation and feature extraction with the ABCD rule. *J Digit Imaging* 33: 574–585. <https://doi.org/10.1007/s10278-019-00316-x>
44. Tsafack N, Kengne J, Abd-El-Atty B, Iliyasu AM, Hirota K, Abd EL-Latif AA (2020) Design and implementation of a simple dynamical 4-D chaotic circuit with applications in image encryption. *Inf Sci* 515:191–217. <https://doi.org/10.1016/j.ins.2019.10.070>
45. Tschandl P, Codella N, Akay NB, Argenziano G, Braun PR, Cabo H, Gutman D, Halpern A, Helba B, Wellenhold RH, Lallas A (2019) Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study. *Lancet Oncol* 20(7):938–947. [https://doi.org/10.1016/S1470-2045\(19\)30333-X](https://doi.org/10.1016/S1470-2045(19)30333-X)
46. Vasconcelos CN, Vasconcelos BN (2017) Experiments using deep learning for dermoscopy image analysis. *Pattern Recogn Lett*. <https://doi.org/10.1016/j.patrec.2017.11.005>
47. Wang Q, Rabab KW (2007) Fast image/video contrast enhancement based on weighted thresholded histogram equalization. *IEEE Trans Consumer Electron* 53(2):757–764. <https://doi.org/10.1109/TCE.2007.381756>
48. Yoshida T, Celebi ME, Schaefer G, Iyatomi H (2016) Simple and effective pre-processing for automated melanoma discrimination based on cytological findings. In: IEEE International Conference on Big Data (Big Data). IEEE, Washington, 3439–3442. <https://doi.org/10.1109/BigData.2016.7841005>

49. Yu L, Chen H, Dou Q, Qin J, Heng PA (2016) Automated melanoma recognition in dermoscopy images via very deep residual networks. *IEEE Trans Med Imaging* 36(4):994–1004. <https://doi.org/10.1109/TMI.2016.2642839>
50. Yüksel ME, Borlu M (2009) Accurate segmentation of dermoscopic images by image thresholding based on type-2 fuzzy logic. *IEEE Trans Fuzzy Syst* 17(4):976–982. <https://doi.org/10.1109/TFUZZ.2009.2018300>
51. Zachary HR, Secrest AM (2019) Public health implications of google searches for sunscreen, sunburn, skin cancer, and melanoma in the United States. *Am J Health Promot* 33(4):611–615. <https://doi.org/10.1177/0890117118811754>
52. Zhang X, Zhao S (2018) Segmentation preprocessing and deep learning based classification of skin lesions. *J Med Imaging Health Inform* 8(7):1408–1414. <https://doi.org/10.1166/jmihi.2018.2448>
53. Zhang J, Xie Y, Xia Y, Shen C (2019) Attention residual learning for skin lesion classification. *IEEE Trans Med Imaging* 38(9):2092–2103. <https://doi.org/10.1109/TMI.2019.2893944>
54. Zhang C, Wu X, Gao X (2019) An improved Gaussian mixture modeling algorithm combining foreground matching and short-term stability measure for motion detection. *Multimed Tools Appl*, 1–23. <https://doi.org/10.1007/s11042-019-08210-y>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.