Ferdowsi University of Mashhad

Journal of Computer and Knowledge Engineering

https://cke.um.ac.ir

Information and Communication Technology Association of Iran

Diagnosing of Skin Lesions Using Deep Convolutional Neural Network and Support Vector Machines*

Research Article

Seyede Tara Naghshbandi[1] ,Abdolhossein Fathi[2]

DOI: 10.22067/cke.2024.83294.1088

Abstract: The number of fatalities resulting from skin cancer has significantly increased over the past few years. Early diagnosis is highly important for the quick treatment of skin cancer. Computer-based dermoscopy analysis methods provide considerable information about the lesions that can be helpful to skin experts in the early detection of skin lesions. These computer-based diagnostic systems require image-processing algorithms to provide mathematical explanations of suspicious areas. Convolutional Neural Network (CNN) as one of the deep learning algorithms has high scalability in interaction with big data, and can automatically extract key image features for classification and segmentation of images. In this study, a hybrid model consisting of deep learning and machine learning method is proposed to classify different types of skin lesions. In this model, at first, an input image is pre-processed to remove the negative effect of Hairs on skin lesion detection and also to prepare it for applying to an efficient deep convolutional network employed as a feature extractor. Then Support Vector Machine (SVM) is utilized as a classifier to detect and classify different types of skin lesions.

Keywords: Skin Cancer Detection, Skin Lesions Classification, Convolutional Neural Network, Support Vector Machine.

1. Introduction

SKIN cancer is the most common type of cancer. Like other cancers, skin cancers, initially present with non-cancerous lesions. These lesions are spots on the skin that are not cancerous but can develop into cancer over time. There are two main types of skin cancer: melanoma and non-melanoma. Although only one percent of all skin cancers are melanoma, it is the most dangerous type of skin cancer due to its tendency to metastasize. According to the Skin Cancer Foundation, The number of melanoma deaths is expected to increase by 6.5 percent in 2022, and also In the past decade (2012 – 2022), the number of new invasive melanoma cases diagnosed annually increased by 31 percent. [1]. Detecting melanoma in the early stages significantly increases the patient's chances of survival. Therefore, to deal with the increase in melanoma mortality, experts and advanced equipment are necessary for accurate and timely diagnosis. Computer-based dermoscopy analysis methods provide considerable information about the lesions that can be helpful to skin experts in the early detection of skin lesions [2]. Dermoscopy is a specialized method for imaging the skin with high resolution, which reduces the reflection of the skin surface and allows doctors to see deeper underlying structures. In recent years, deep learning algorithms have entered virtually all technology fields and have achieved good results in different fields such as biometrics, self-driving car, robotics, photo description, earthquake prediction, and medical image analysis. Nowadays, various deep learning models have become important in the field of medical image recognition due to their ability to automatically recognize image patterns. Since the convolutional neural network has shown significant performance in recognizing patterns in digital images, the focus of this study is to use it to extract key features from skin lesions.

2. Related Work

In recent years several deep learning-based models were employed for skin lesion detection and lassification. In a study conducted by Gessert et al. [3], The EfficientNet architecture, which is one of the convolutional neural network architectures, was used for skin lesions detection. They also utilized DenseNet layers as a classifier in the final model. A new optimized technique was proposed for skin cancer diagnosis from the input images by Zhang et al. The method was based on a Convolutional neural network. An improved version of the whale optimization algorithm was adopted for optimizing the efficiency result of CNN. The utilized optimization algorithm was adopted for the optimal selection of weights and biases in the network to minimize the error of the network output and the desired output [4]. Maron et al. proposed a convolutional neural network to classify skin lesions into five categories [5]. Khouloud et al. developed two deep learning models for an effective

* Manuscript received: 2023 July 7, Revised, 2024 May 13, Accepted, 2024 May 26.

[1] MSc. Student at Department of Computer Engineering, Razi University, Kermanshah, Iran.

[2] Corresponding author. Associate Professor at Department of Computer Engineering, Razi University, Kermanshah, Iran.

Email: a.fathi@razi.ac.ir.

segmentation-based classification framework. The operational principle of the presented model includes a set of three phases namely pre-processing, attention W-net-based segmentation, and Inception Resnet block-based lesion classification. The dual encoder-decoder architectures (W-net) which are composed of a ResNet EncoderDecoder, a ConvNet Encoder-Decoder, and a Feature Pyramid network have been developed to create a robust CNN architecture to provide solutions to skin lesion segmentation issues [6]. In another study conducted by Adegun et al. [7], a deep convolutional network-based architecture was proposed for robust detection and segmentation of melanoma lesions. This architecture adopts an enhanced deep convolutional network interconnected with a series of skip pathways. It also employs a reduced-size encoder-decoder network that minimizes computational resource consumption. They used the ISIC2017 dataset in their study which contains dermoscopy images with different image sizes with the highest resolution of 1022 x 767. 2000 dermoscopy images and 600 dermoscopy images were used for training and testing respectively. Hoshyar et al. represented the pre-processing techniques required for designing the automatic skin cancer detection system. They classified the whole process into two sections, Image Enhancement, and Image Restoration. In these two processes, all the steps with their beneficial techniques to enhance the skin cancer images and also the useful filters to remove the noise and smooth the images were explained. In another study by Yilmaz et al. [9], three convolutional neural network architectures were used to diagnose benign or malignant skin lesions. These architectures included Alexnet, GoogleNet, and ResNet-50. Qin et al. [10] used a generative adversarial network (GAN) to improve the diagnosis of skin lesions. GAN neural networks are generative models that generate new data similar to training data. In this study, first, artificial data were generated using the GAN algorithm. Then images were categorized by employing the transfer learning technique on the pre-trained ResNet-50 network. Annaby et al. [11] introduced a graph signal processing approach to detect malignant melanoma via dermoscopic images on the ISIC2017 dataset. They transformed each dermoscopic image into a graph of superpixel vertices. The graph signals, as well as weighted edges, were defined using color, geometric and texture features. Several graph features were extracted in both time and frequency domains. Gong et al. [12] used generative adversarial networks (GANs) to create a balanced sample space to better train convolutional neural networks (CNNs) on the ISIC2019 dataset. Through transfer learning, the pre-training CNNs were used for finetuning, then the effects of different CNNs on the classification of different categories of dermoscopy images were compared, and the CNNs with better classification effects were selected for the ensemble of different strategies. Masni et al. [13] after balancing the data using the data augmentation method, categorized the images using different convolutional neural network architectures such as Inception-v3, ResNet-50, Inception-ResNet-v2, and DenseNet-201. Hameed et al. [14] used Alexnet architecture and classified skin images into four categories: Healthy, Eczema, Benign, and Malignant. Kassem et al. [15] tried to classify skin images into 8 classes using the GoogleNet architecture. They tried to fine-tune the employed model by replacing the last two layers of GoogleNet architecture with multiclass SVM. Albahar [16] proposed a new prediction model that classifies skin lesions into benign or malignant lesions based on a novel regularizer technique. It was a binary classifier that discriminates between benign or malignant lesions using the ISIC2017 dataset. Dinga et al. [17] employed a deep attention branch network (DABN) model along with an entropy-guided loss weighting (ELW) strategy to detect and classify skin lesions. Alsaade et al. [18] also used the image segmentation method and deep learning models. To this end, the segmented images were applied to the Alexnet and ResNet-50 architectures. Hsu and Tseng [19] proposed hierarchy-aware contrastive learning with late fusion (HAC-LF), to perform multi-class skin lesions classification. The late fusion method was applied to balance the major-type and multi-class classification performance. Wei et al. [20] first used noise reduction methods to remove image noise from input images, then they used image segmentation to segregate the cancerous lesion from them. After extracting the features from the segmented images, they selected the key features of the image using the thermal exchange optimization algorithm. Finally, they gave the selected features as input to the SVM algorithm to perform the classification task. Fernando et al. presented a self-adapting weighting approach and introduced a novel loss function, named DWB loss. The weighting scheme was based on the class frequency of training data and the prediction difficulty of individual data instances. The prediction difficulty was determined by the prediction score produced by the neural network [21]. Pacheco et al. proposed the Metadata Processing Block (MetaBlock), an attention-based mechanism approach that used the metadata to enhance the feature maps extracted from images to improve data classification [22]. In another study, Sun et al. proposed a single baseline for skin lesion classification which used the information of data augmentation as additional patient information. The metadata used in their manuscript included additional info that was generated during data augmentation [23]. Putra et al. proposed a novel technique to perform dynamic pre-processing on the inference (DPI). They used two models to accomplish their goal. The first model was for predicting the skin condition, and the second model was built to predict the best augmentation for the inference stage (PA). They used EfficientNet b4 for both models [24]. Banerjee et al. Proposed ta deep learning-based 'Keras' algorithm, which was established on the implementation of DCNNs to investigate melanoma from dermoscopic and digital pictures and provided swifter and more accurate results as contrasted to standard CNNs [25]. In another study, Vaiyapuri et al. developed a novel computational intelligence-based melanoma detection and classification technique using dermoscopic images (CIMDC-DIs). The proposed CIMDC-DI model encompassed different subprocesses. Primarily, bilateral filtering with fuzzy k-means (FKM) clustering-based image segmentation was applied as a pre-processing step. Besides, a NasNet-based feature extractor with stochastic gradient descent was applied for feature extraction. Finally, the manta ray foraging optimization

(MRFO) algorithm with a cascaded neural network (CNN) was exploited for the classification process [26]. Kaur et al. used a new convolutional neural network to classify versions of the ISIC dataset. They called their proposal method lesion classification network (LCNet). Their proposed model, inspired by the deep convolutional neural network for skin cancer classification, was trained in an end-to-end manner on dermoscopic skin cancer images. Three different datasets from the ISIC challenge were incorporated to perform the experiments, and an additional PH2 set was used for testing [27]. Another study by Chabi Adjobo et al. proposed a data augmentation method based on multiscale image decomposition and pulse coupled neural network (PCNN) fusion strategy as a solution to alleviate the lack of labeled dermoscopic data in all existing skin tones [28]. Adegun et al. proposed a new framework that performs both segmentation and classification of skin lesions for automated detection of skin cancer. Their proposed framework consists of two stages: the first stage leverages on an encoder-decoder Fully Convolutional Network (FCN) to learn the complex and inhomogeneous skin lesion features with the encoder stage learning the coarse appearance and the decoder learning the lesion borders details [29]. In another study three ensemble classification models based on input feature manipulation from the shape properties, color variation, and texture analysis were presented: the SE-OPF, SEFSOPF, and FEFS OPF algorithms [30]. Mahbod et al. investigated the impact of using skin lesion segmentation masks on the performance of dermatoscopic image classification. To do this, first, they developed a baseline classifier as the reference model without using any segmentation masks. Then, they used either manually or automatically created segmentation masks in both training and test phases in different scenarios and investigated the classification performances [31]. in a comprehensive study, Haggenmüller et al. investigated the current state of research on melanoma and assessed their potential clinical relevance by evaluating three main aspects: test set characteristics, test setting (experimental/clinical, the inclusion of metadata), and representativeness of participating clinicians [32].

Höhn et al. investigated whether a combination of histologic whole slides image (WSI) analysis based on convolutional neural networks (CNNs) and commonly available patient data (age, sex, and anatomical site of the lesion) in a binary melanoma/nevus classification task could increase the performance compared with CNNs alone [33]. Dhivyaa et al. proposed a hybrid approach by using decision trees and random forest algorithms. According to their claims, compared to the other machine learning algorithms like neural networks and support vector machines, the proposed model is much simpler with less computation complexity and more accuracy in the classification of skin lesion images [34]. Thanh et al. proposed a method to detect melanoma skin cancer with automatic image processing techniques. Their method includes three stages: pre-process images of skin lesions by adaptive principal curvature, segment skin lesions by the color normalization, and extract features by the ABCD rule [35]. Song et al designed a novel end-to-end multi-task framework that could classify, detect and segment dermoscopy images simultaneously for skin lesion analysis. The framework could take the image with arbitrary size as the input and outputs melanoma type, position, and boundary without any additional post-processing operations [36]. Thornhofer-Hamsey et al. performed binary and multi-class classification using the transfer learning method. They first augmented the images in the HAM10000 dataset using the data augmentation method and then gave these images as input to the densenet201 architecture [37]. A Hybrid Approach using the Fuzzy Logic System and the Modified Genetic Algorithm was proposed by Saurabh Jha et al. According to their proposed method the modified genetic algorithm is used to select the best features which will participate in the fuzzy rules generation process. It selects the best features along with the calculation of the accuracy of the system based on the selected features. A new rule reduction algorithm (RR_algorithm) is then utilized to reduce the certain number of rules to decrease the complexity of the rule base of the fuzzy system [38]. Another study by Araújo et al. used a melanoma segmentation method based on U-net and LinkNet deep learning networks combined with transfer learning and fine-tuning techniques. The experiments were carried out in three datasets including PH2, ISIC 2018, and DermIS [39].

Shorfuzzaman et al. introduced an explainable CNN-based deep learning stacked ensemble framework using the transfer learning concept for melanoma skin cancer detection. Their study sought to attain this by developing an ensemble network where prediction results from multiple CNN sub-models were combined and fed to a meta-learner for the final prediction of malignant melanoma moles [40]. The application of deep learning is not limited to the detection of skin lesions, for example in [41] a deep learning method for the classification and detection of metastatic cancer in histopathologic images of lymph node sections was proposed. A diagnostic method of cancer in histopathologic images is time-consuming and tedious for pathologists because a large tissue area has been examined, and tiny metastasis can be easily ignored. Thus the developed deep learning method can help pathologists in examining the histopathologic scans and assist in decision-making to analyze the disease and cancer staging, which will give consequential opinions in clinical diagnosis.

Despite the strengths of deep learning algorithms, their need for large datasets for training the huge parameters of employed models and also the need for powerful hardware to train the network parameters caused to employ the pre-trained models in the literature. Since the similarity of different classes of skin lesions is high, to achieve the desired result, it is necessary to determine the parameters and arrangement of the network architecture precisely. Therefore, employing a pre-trained model may not classify the skin lesion properly. To this end, a combination of a convolutional neural network (which recent studies show its superiority to other models for skin lesion detection) and a support vector machine (as one of the powerful machine learning-based classifiers) is employed to detect and categorize different types of skin lesions.

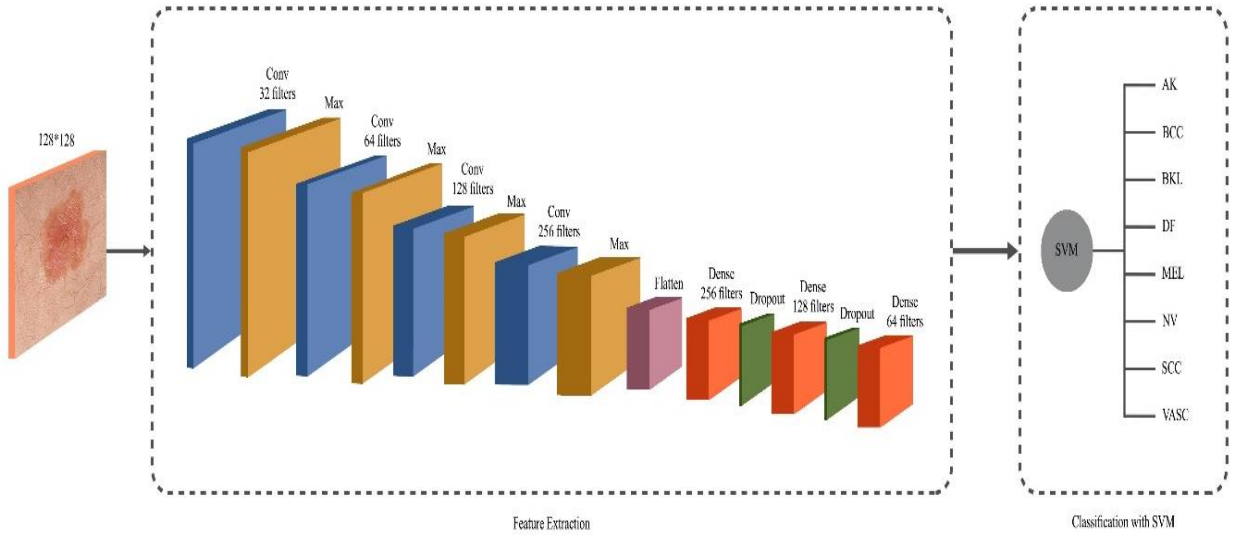

Figure 1. The architecture of the proposed method using CNN+SVM

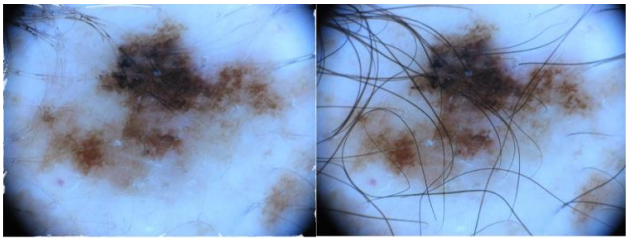

Figure 2. A view of a skin lesion in which the excess hair in the image has been removed

3. Methodology

This paper proposed a hybrid method to detect and categorize skin lesion images. The goal is to categorize different types of skin lesions in the ISIC2019 dataset including Melanoma (MEL), Melanocytic Nevus (NV), Basal Cell Carcinoma (BCC), Actinic Keratosis (AK), Benign Keratosis (BKL), Dermatofibroma (DF), Vascular lesion (VASC), and Squamous Cell Carcinoma (SCC) and also ISIC2020 which is summarized into two categories, Benign and Malignant. To this end, a combination of convolutional neural networks (CNN) and support vector machines (SVM) is employed. The CNN is used as a feature extractor and SVM is utilized as a classifier. The structure of layers and parameters are modified to have high efficiency in classifying skin lesions. The architecture of the proposed method is shown in Figure 1.

3.1. Image pre-processing

The purpose of preprocessing step is to prepare the data for neural network entry. At first, the numerical range of image pixels is normalized, because neural networks tend to converge towards larger numbers. Since the value of each pixel is a number between 0 and 255, it is necessary to reduce this range between zero and one. Also, since the dimensions of input images may not the same, it is necessary to make the dimensions of all images the same, in this study the size of input images is resized to 128×128 pixels. Another pre-processing step in this study is the removal of hair from the images because in some images, a person's body hair is on the lesion and this could make lesion detection more difficult. For this purpose, a method called Dullrazor [42] is used for removing hair from the images. Figure 2 shows an image of a skin lesion that has been well removed by this technique.

3.2. Convolutional Neural Network

After applying preprocessing step, a convolutional neural network with 15 layers is used as a feature extraction module. The specifications of these layers and the number of each type are given in Table 1.

Table 1. CNN network specifications

Layer Name	Number of Layers
Input_Layer	1
Conv2D_layer	4
Maxpooling2D_Layer	4
Dropout_Layer	2
Flatten_Layer	1
Dense_Layer	3

The point to be noted is that the high similarity between skin lesions has made it difficult even for a physician to detect its type. Therefore, based on the similarity of skin lesion classes to each other, it is necessary to select a small

kernel size to scroll the image pixels and apply convolution multiplication operations on a smaller scale. To this end, the kernel size parameter of convolution layers is set to 2. After convolutional layers, the Max-pooling layer is used to halve the size of the image and reduce the computational overhead of the network. The flatten layer is then used. This layer flattens the multi-dimensional input tensors into a single dimension, so we can pass the data into every single neuron of the model effectively. Then three fully connected layers, which are also known as the dense layers, are used. This layer is a simple layer of neurons in which each neuron receives input from all the neurons of the previous layer. After each dense layer, a dropout layer is used. In the training phase, this layer resets the output of some neurons to zero, depending on the probability value assigned to it. In this case, we are faced with a network that has to extract important features in the face of data without the use of disabled neurons; this will prevent the network from overfitting.

The proposed CNN is used as a feature extractor and SVM with RBF kernel is used as a classifier.

The SVM has several kernels, the most important of which are linear, polynomial, RBF, and sigmoid. Kernels are a set of mathematical functions that take data as input and transform it into the required form. In other words, kernel functions are used to map the original features into a higher dimensional space to make a linear decision on it.

4. Materials

In this section, we describe the materials and our experiments.

4.1. Datasets

In our study, the ISIC2019 and ISIC2020 datasets [44,45] were used for the evaluation of the proposed method. ISIC2019 dataset includes 25331 dermoscopic images, collected by the International Skin Imaging Collaboration (ISIC). These images were manually categorized into 8 classes that represent eight types of skin lesions. The number of images in each is shown in Table 2. An example of these classes is shown in Figure 3.

Table 2. Details of ISIC2019 dataset

Class Name	Class Full Name	Number of Samples
MEL	Melanoma	4,522
NV	Melanocytic nevus	12,857
BCC	Basal cell carcinoma	3,323
AK	Actinic keratosis	867
BKL	Benign keratosis	2,624
DF	Dermatofibroma	239
VASC	Vascular lesion	253
SCC	Squamous cell carcinoma	628

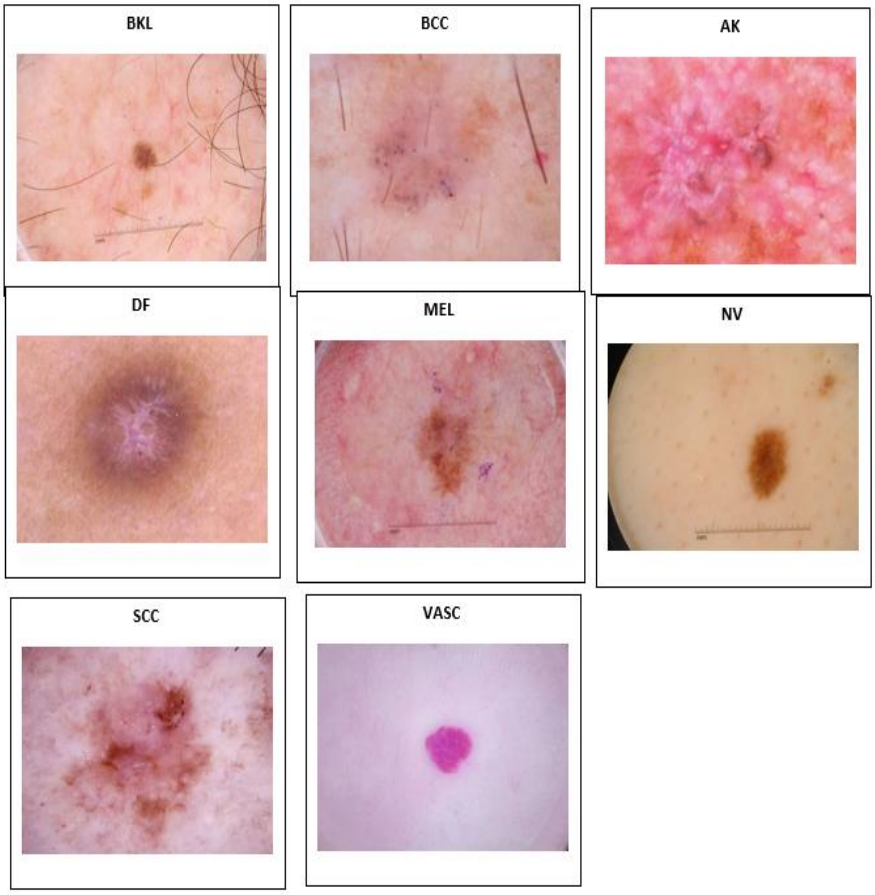

Figure 3. An image of ISIC 2019 classes

Figure 4. Applying horizontal flip on the image using the data augmentation technique

This dataset consists of two datasets called HAM10000, which contains dermoscopic images with a size of 450×600, and the BCN-20000 dataset, which includes dermoscopic images with the size of 1024×1024. The combination of data from these two datasets has formed the ISIC2019 dataset. Unlike the ISIC2019 dataset, the ISIC2020 dataset is used for binary classification tasks. This dataset has 33126 images which are divided into benign and malignant classes. The number of images in each is shown in Table 3.

Table 3. Details of the ISIC2020 dataset

Class Name	Number of Samples
Benign	32542
Malignant	584

Table 4. Details of augmented ISIC2019 dataset.

Class Name	Number of training samples	Number of validation samples
MEL	11482	4922
NV	10969	4702
BCC	12761	5469
AK	10647	4563
BKL	11245	4820
DF	10857	4653
VASC	11033	4729
SCC	10449	4479

Table 5. Details of augmented ISIC2020 dataset.

Class Name	Number of training samples	Number of validation samples
Benign	22779	9763
Malignant	23302	9987

4.2. Data Augmentation

Due to unbalanced classes, it is necessary to use a technique called data augmentation to be able to balance the number of images in different classes of the ISIC2019 and ISIC2020 datasets. By using the data augmentation technique, it is possible to create artificial images that are similar to their original prototype in all respects, more information about this technique is available in [43]. In this method, a set of images is given to the algorithm, and in the next step, by applying a set of operations such as rotation, Width Shifting, Height Shifting, Brightness, Zoom, Horizontal Flip, and Vertical Flip, new images can be created. For example, in figure 4 the horizontal flip operation is applied to an image of a skin lesion.

Using the data augmentation technique, 102449 new artificial images were created and added to the ISIC2019 dataset to have 127780 images. For the ISIC2020 dataset, we did not change the number of benign class images, and just due to the rarity of malignant class images in the ISIC2020 dataset, 32704 artificial images of this class type were produced by applying the data augmentation technique. The number of images for all classes for both datasets are almost the same and is shown in Table 4 and 5.

4.3. Model parameters optimization

After creating the proposed model, it is necessary to determine the network hyperparameters. The most important parameters are the optimizer function and learning rate, as well as the loss function. Optimizers are algorithms or methods used for changing the attributes of neural networks such as weights and learning rate to reduce the loss rate. For this purpose, the Adam optimizer function is used. This function is one of the best optimizer functions because it requires little memory space and works well with large datasets and large parameters. The learning rate indicates the speed of updating the weights. This hyper-parameter should neither be too large (because it could override the optimal state) nor be too small (because the network convergence takes too long). Usually, the value of the learning rate is determined by trial and error. In this study, the value of 0.001 was set for the learning rate of the Adam function. Sometimes during training, the value of the loss function does not go beyond a certain value and remains constant and stuck in a local minimum. In such cases, it is necessary to use a parameter called decay. This parameter reduces the learning rate value in each epoch; thus, it prevents the loss function from being caught in local minima. The optimum value for this parameter is equal to dividing the value of the learning rate by the number of epochs. In this study, the value of epoch was set equal to 100, so for the decay parameter, the value of 0.00001 is obtained. The last hyper-parameter to be

determined is loss function. This function is used for determining whether the algorithm is working properly. The loss function is a function that calculates the distance between the current output of the algorithm and the expected output and is a way to evaluate how the algorithm models the data. Since the used dataset in this study has more than two classes, the categorical cross-entropy loss function is used. In the following section, the results of the proposed method, as well as the evaluation metrics, are examined.

5. Model evaluation and comparison

To evaluate the performance of the proposed method, 3-fold cross-validation technique was used.

After training the proposed model in each fold, the performance of the proposed method was measured on the validation set in terms of different metrics such as Kappa, Precision, Recall, F1-measure, Accuracy, and Balanced Accuracy.

One of the most important cases of using the kappa coefficient is in diagnosing and testing processes in medical topics. Kappa measures the agreement between two raters who each classify N items into C mutually exclusive categories, which was later used by the machine-learning community to measure the performance of classification algorithms. This metric is defined by the following equation:

$$Kappa = \frac{(po-pe)}{(1-pe)} = 1 - \frac{1-po}{1-pe} \quad (1)$$

where po is the relative observed agreement among raters, and pe is the hypothetical probability of chance agreement, using the observed data to calculate the probabilities of each observer randomly seeing each category. Kappa value between 0.4 and 0.6, indicates the average performance of the classification algorithm, and values between 0.6 to 0.8 and 0.8 to 1 indicate the good and excellent performance of the algorithm, respectively.

Precision is defined as the number of true positives (TP) over the number of true positives plus the number of false positives (FP). It is defined by the following equation:

$$\text{Precision} = \frac{\text{TP}}{TP + FP} \quad (2)$$

Recall is defined as the number of true positives (TP) over the number of true positives plus the number of false negatives (FN). It is defined by the following equation:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

F1-measure is defined as the harmonic mean of precision and recall. Its equation is defined by:

$$F1 - Measure = \frac{2\times\text{Recall}\times\text{Precision}}{\text{Recall}+\text{Precision}} \quad (4)$$

Accuracy is the most intuitive performance measure and it is simply a ratio of correctly predicted observations to the total observations but it is not an appropriate metric when the dataset is unbalanced and we should use other metrics like balanced accuracy and f1-measure. The equation of accuracy metric is defined by:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+\text{FP}+FN} \quad (5)$$

Balanced Accuracy is the arithmetic mean of sensitivity and specificity; its use case is when dealing with unbalanced data when one of the target classes appears a lot more than the other. The equation of this metric is defined by:

$$\text{Balanced Accuracy} = \frac{\text{Recall} + \text{Precision}}{2} \quad (6)$$

The obtained precision, Recall, and F1-measure results of the proposed method on each class and for each fold are shown in Tables 6 and 7 based on the used datasets. Also, the average of all metrics in each fold is shown in this Table.

For more comparison, the obtained results of the proposed method along with the results of related works that performed multi-class classification are also shown in Tables 8 and 9. It should be noted that in some papers, the type of classification is not mentioned and for a fair comparison, only the papers that have done multi-class classification are shown in the table.

Since different metrics are used in each research to evaluate the model, and some studies have only provided the total accuracy of their method and have not published the accuracy for each class, for better comparison, the overall accuracy of the methods is compared. Also, the effect of employing SVM as a classifier is investigated. In this case, the classification of different classes is done by replacing SVM with a dense layer along with softmax activation at the end of the proposed CNN model. This case is named "Proposed Method (CNN-SVM)" in Tables 8 and 9. From the obtained results, it is clear that the proposed method has the highest Balance Accuracy. Also, the precision, Recall, F1-measure, and Kappa of the proposed method are higher than the other existing methods.

Table 6. Results of the proposed method on the ISIC2019 dataset (CNN+SVM)

Fold1			
Class Name	Precision	Recall	F1-measure
AK	0.7968	0.8304	0.8133
BCC	0.7310	0.7731	0.7514
DF	0.6696	0.6376	0.6532
SCC	0.9524	0.9708	0.9615
VASC	0.7292	0.6325	0.6774
MEL	0.7814	0.8111	0.7960
BKL	0.8374	0.8486	0.8430
NV	0.9830	0.9930	0.9880
Total Accuracy=0.8094			
Total Balance Accuracy =0.8121			
Total AUC=0.8924			
Total F1-measure=0.8104			
Total Recall=0.8121			
Total Precision=0.8101			
Kappa= 0.7820			
Fold2			
AK	0.8435	0.8538	0.8486
BCC	0.7477	0.8387	0.7906
DF	0.7408	0.7120	0.7261
SCC	0.9709	0.9691	0.9700
VASC	0.7924	0.7020	0.7445
MEL	0.8496	0.8135	0.8312
BKL	0.8493	0.8908	0.8896
NV	0.9909	0.9945	0.9927
Total Accuracy=0.8450			
Total Balance Accuracy = 0.8467			
Total AUC=0.9122			
Total F1-measure=0.8466			
Total Recall=0.8468			
Total Precision=0.8491			
Kappa= 0.8227			
Fold3			
AK	0.8242	0.8087	0.8164
BCC	0.7175	0.7883	0.7512
DF	0.6935	0.6531	0.6727
SCC	0.9484	0.9759	0.9620
VASC	0.7152	0.6786	0.6964
MEL	0.7728	0.7929	0.7827
BKL	0.8683	0.8127	0.8396
NV	0.9735	0.9962	0.9847
Total Accuracy=0.8348			
Total Balance Accuracy = 0.8132			
Total AUC=0.8931			
Total F1-measure=0.8132			
Total Recall=0.8132			
Total Precision=0.8141			
Kappa= 0.7843			

Table 7. Results of the proposed method on the ISIC2020 dataset (CNN+SVM)

Fold1			
Class Name	Precision	Recall	F1-measure
Benign	0. 9771	0. 9689	0. 9729
Malignant	0. 9698	0. 9778	0. 9738
Total Accuracy=0.9734			
Total Balance Accuracy =0.9733			
Total AUC=0.9733			
Total F1-measure=0.9734			
Total Recall=0.9733			
Total Precision=0.9734			
Kappa= 0.9467			
Fold2			
Benign	0. 9808	0. 9753	0. 9781
Malignant	0. 9760	0. 9814	0. 9814
Total Accuracy=0. 9784			
Total Balance Accuracy = 0.9783			
Total AUC=0. 9783			
Total F1-measure=0. 9784			
Total Recall=0. 9783			
Total Precision=0. 9784			
Kappa=0.9567			
Fold3			
Benign	9.9783	0.9682	0.9732
Malignant	0.9693	0.9790	0.941
Total Accuracy=0.9736			
Total Balance Accuracy=0.9736			
Total AUC=0.9736			
Total F1-measure=0.9737			
Total Recall=0.9736			
Total Precision=0.9738			
Kappa= 0.9473			

6. Conclusions

Early detection of cancer is highly serious because the chances of treatment are very high in the early stages. Skin cancer is the most dangerous type of cancer that has affected the world's population. Nowadays, Computer-aided diagnostic systems are growing rapidly and a variety of algorithms are used to detect cancer lesions, each of which has advantages and disadvantages. Unlike other related works that have mainly used transfer-learning methods, we used a customized convolutional neural network to improve the efficiency and performance of skin lesions detection and classification. In the proposed method, we used CNN and SVM as a hybrid approach. The former was used as a feature extractor and the latter as a classifier, which was able to detect and classify more than 84% and 97% of skin lesions in the ISIC2019 and ISIC2020 datasets. Since deep learning algorithms require many data to train the network to achieve the desired performance, the lack of access to training data is one of the biggest challenges in this field. To this end, the proposed method used data augmentation techniques for producing artificial images. Although this technique solves this problem to some extent, due to the lack of data diversity, the network may not be able to correctly identify the desired data class when dealing with new data. The obtained results show that the proposed method performed better than state-of-the-art methods. Also obtained results show that the use of combined methods can be an effective solution in classification tasks.

Table 8. Comparison of the proposed method with other previous related works using the ISIC2019 dataset

Source	Method	No. Class	dataset	Accuracy	Precision	Recall	F1	AUC	Kappa	Balance Accuracy
[3]	EfficientNet	8	ISIC 2019	-	-	0.63	-	-	-	-
[12]	Ensemble model using GAN	8	ISIC 2019	0.924	-	0.483	0.488	0.919	-	-
[15]	Deep Neural Network GoogleNet	8	ISIC 2019	0.81	0.77	0.74	-	-	-	-
[19]	HAC-LF	8	ISIC 2019	0.871	-	0.842	-	-	-	-
[21]	DWB loss approach	8	ISIC 2019	-	0.69	0.66	0.67	0.82	-	-
[22]	Attention-based mechanism	8	ISIC 2019	0.913	-	-	-	0.865	-	0.522
[23]	Additional patient information method	8	ISIC 2019	-	-	0.662	-	0.915	-	0.662
[24]	PA+DPA	8	ISIC 2019	0.950	-	0.650	0.640	0.910	-	0.650
Proposed	Method (CNN-SVM)	8	ISIC 2019	0.8450	0.8491	0.8468	0.8467	0.9122	0.7938	0.8467

Table 9. Comparison of the proposed method with other previous related works using the ISIC2020 dataset

Source	Method	No. Class	dataset	Accuracy	Precision	Recall	F1	AUC	Kappa	Balance Accuracy
[25]	Keras model	2	ISIC 2020	0.970	0.913	0.960	-	-	-	-
[26]	CIMDC-DI	2	ISIC 2020	0.960	0.960	0.959	0.959	0.891 0.902 0.917	-	-
[27]	lesion classification network (LCNet)	2	ISIC 2020	0.9042	0.9048	0.9039	0.9041	-	-	-
[28]	PCNN	2	ISIC 2020	0.955	0.936	0.965	0.950	-	-	-
Proposed	Method (CNN-SVM)	2	ISIC 2020	0.975133	0. 9752	0. 975067	0. 975167	0. 97506	0.950233	0.97507

7. Acknowledgment

Declarations:

Funding (information that explains whether and by whom the research was supported): No funds, grants, or other support was received.

Conflicts of interest/Competing interests (include appropriate disclosures): The authors have no conflicts of interest to declare that are relevant to the content of this article.

Availability of data and material: The data are available at https://challenge2019.isic-archive.com/data.html and https://challenge2020.isic-archive.com/

Authors' contributions: All authors contributed to the study's conception and design. Material preparation, data collection, analysis, and writing of the first draft of the paper were performed by Seyede Tara Naghshbandi. The supervision and writing, review, and editing of the manuscript were done by Abdolhossein Fathi. All authors read and approved the final manuscript.

8. References

[1] https://www.skincancer.org/skin-cancer-information/skin-cancer-facts/

[2] Calderón, C., Sanchez, K., Castillo, S., & Arguello, H. (2021) BILSK: A bilinear convolutional neural network approach for skin lesion classification. Computer Methods and Programs in Biomedicine Update, 1, 100036.

[3] Gessert, N., Nielsen, M., Shaikh, M., Werner, R., & Schlaefer, A. (2020). Skin lesion classification using ensembles of multi-resolution EfficientNets with metadata. MethodsX, 7, 100864.

[4] Zhang, N., Cai, Y. X., Wang, Y. Y., Tian, Y. T., Wang, X. L., & Badami, B. (2020). Skin cancer diagnosis based on optimized convolutional neural network. Artificial intelligence in medicine, 102, 101756.

[5] Maron, R. C., Weichenthal, M., Utikal, J. S., Hekler, A., Berking, C., Hauschild, A., ... & Thiem, A. (2019). Systematic outperformance of 112 dermatologists in multiclass skin cancer image classification by convolutional neural networks. European Journal of Cancer, 119, 57-65.

[6] Khouloud, S., Ahlem, M., Fadel, T., & Amel, S. (2021). W-net and inception residual network for skin lesion segmentation and classification. Applied Intelligence, 1-19.

[7] Adegun, A. A., & Viriri, S. (2019). Deep learning-based system for automatic melanoma detection. IEEE Access, 8, 7160-7172.

[8] Hoshyar, A. N., Al-Jumaily, A., & Hoshyar, A. N. (2014). The beneficial techniques in preprocessing step of the skin cancer detection system comparing. Procedia Computer Science, 42, 25-31.

[9] Yilmaz, E., & Trocan, M. (2020, March). Benign and malignant skin lesion classification comparison for three deep-learning architectures. In Asian conference on intelligent information and database systems (pp. 514-524). Springer, Cham.

[10] Qin, Z., Liu, Z., Zhu, P., & Xue, Y. (2020). A GAN-based image synthesis method for skin lesion classification. Computer Methods and Programs in Biomedicine, 195, 105568.

[11] Annaby, M. H., Elwer, A. M., Rushdi, M. A., & Rasmy, M. E. (2021). Melanoma detection using spatial and spectral analysis on superpixel graphs. Journal of digital imaging, 34(1), 162-181.

[12] Gong, A., Yao, X., & Lin, W. (2020). Classification for dermoscopy images using convolutional neural networks based on the ensemble of individual advantage and group decision. IEEE Access, 8, 155337-155351.

[13] Al-Masni, M. A., Kim, D. H., & Kim, T. S. (2020). Multiple skin lesions diagnostics via integrated deep convolutional networks for segmentation and classification. Computer methods and programs in biomedicine, 190, 105351.

[14] Hameed, N., Shabut, A. M., Ghosh, M. K., & Hossain, M. A. (2020). Multi-class multi-level classification algorithm for skin lesions classification using machine learning techniques. Expert Systems with Applications, 141, 112961.

[15] Kassem, M. A., Hosny, K. M., & Fouad, M. M. (2020). Skin lesions classification into eight classes for ISIC 2019 using deep convolutional neural network and transfer learning. IEEE Access, 8, 114822-114832.

[16] Albahar, M. A. (2019). Skin lesion classification using convolutional neural network with novel regularizer. IEEE Access, 7, 38306-38313.

[17] Dinga, S., Wu, Z., Zheng, Y., Liu, Z., Yang, X., Yang, X., Yuan, G., & Xie, J. (2021). Deep attention branch networks for skin lesion classification. Computer Methods and Programs in Biomedicine, 212, 106447.

[18] Alsaade, F. W., Aldhyani, T. H., & Al-Adhaileh, M. H. (2021). Developing a Recognition System for Diagnosing Melanoma Skin Lesions Using Artificial Intelligence Algorithms. Computational and Mathematical Methods in Medicine, 2021.

[19] Hsu, B. W., & Tseng, V.S. (2022). Hierarchy-Aware Contrastive Learning with Late Fusion for Skin Lesion Classification. Computer Methods and Programs in Biomedicine, (Jan. 2022), 106666.

[20] Wei, L., Pan, S. X., Nanehkaran, Y. A., & Rajinikanth, V. (2021). An Optimized Method for Skin Cancer Diagnosis Using Modified Thermal Exchange Optimization Algorithm. Computational and Mathematical Methods in Medicine, 2021.

[21] Fernando, K. R. M., & Tsokos, C. P. (2021). Dynamically weighted balanced loss: class imbalanced learning and confidence calibration of deep neural networks. IEEE Transactions on Neural Networks and Learning Systems.

[22] Pacheco, A. G., & Krohling, R. A. (2021). An attention-based mechanism to combine images and metadata in deep learning models applied to skin cancer classification. IEEE journal of biomedical and health informatics, 25(9), 3554-3563.

[23] Sun, Q., Huang, C., Chen, M., Xu, H., & Yang, Y. (2021). Skin lesion classification using additional patient information. BioMed research international, 2021.

[24] Putra, T. A., Rufaida, S. I., & Leu, J. S. (2020). Enhanced skin condition prediction through machine learning using dynamic training and testing augmentation. IEEE Access, 8, 40536-40546.

[25] Banerjee, S., Singh, S. K., Chakraborty, A., Basu, S., Das, A., & Bag, R. (2021). Diagnosis of Melanoma Lesion Using Neutrosophic and Deep Learning. Traitement du Signal, 38(5).

[26] Vaiyapuri, T., Balaji, P., Alaskar, H., & Sbai, Z. (2022). Computational Intelligence-Based Melanoma Detection and Classification Using Dermoscopic Images. Computational Intelligence and Neuroscience, 2022.

[27] Kaur, R., GholamHosseini, H., Sinha, R., & Lindén, M. (2022). Melanoma Classification Using a Novel Deep Convolutional Neural Network with Dermoscopic Images. Sensors, 22(3), 1134.

[28] Chabi Adjobo, E., Sanda Mahama, A. T., Gouton, P., & Tossa, J. (2022). Towards Accurate Skin Lesion Classification across All Skin Categories Using a PCNN Fusion-Based Data Augmentation Approach. Computers, 11(3), 44.

[29] Adegun, A. A., & Viriri, S. (2020). FCN-based DenseNet framework for automated detection and classification of skin lesions in dermoscopy images. IEEE Access, 8, 150377-150396.

[30] Oliveira, R. B., Pereira, A. S., & Tavares, J. M. R. (2017). Skin lesion computational diagnosis of dermoscopic images: Ensemble models based on input feature manipulation. Computer methods and programs in biomedicine, 149, 43-53.

[31] Mahbod, A., Tschandl, P., Langs, G., Ecker, R., & Ellinger, I. (2020). The effects of skin lesion segmentation on the performance of dermatoscopic

image classification. Computer Methods and Programs in Biomedicine, 197, 105725.

[32] Haggenmüller, S., Maron, R. C., Hekler, A., Utikal, J. S., Barata, C., Barnhill, R. L., ... & Brinker, T. J. (2021). Skin cancer classification via convolutional neural networks: systematic review of studies involving human experts. European Journal of Cancer, 156, 202-216.

[33] Höhn, J., Krieghoff-Henning, E., Jutzi, T. B., von Kalle, C., Utikal, J. S., Meier, F., ... & Brinker, T. J. (2021). Combining CNN-based histologic whole slide image analysis and patient data to improve skin cancer classification. European Journal of Cancer, 149, 94-101.

[34] Dhivyaa, C. R., Sangeetha, K., Balamurugan, M., Amaran, S., Vetriselvi, T., & Johnpaul, P. (2020). Skin lesion classification using decision trees and random forest algorithms. Journal of Ambient Intelligence and Humanized Computing, 1-13.

[35] Thanh, D. N., Prasath, V. B., Hieu, L. M., & Hien, N. N. (2020). Melanoma skin cancer detection method based on adaptive principal curvature, colour normalisation and feature extraction with the ABCD rule. Journal of Digital Imaging, 33(3), 574-585.

[36] Song, L., Lin, J., Wang, Z. J., & Wang, H. (2020). An end-to-end multi-task deep learning framework for skin lesion analysis. IEEE journal of biomedical and health informatics, 24(10), 2912-2921.

[37] Thurnhofer-Hemsi, K., & Domínguez, E. (2021). A convolutional neural network framework for accurate skin cancer detection. Neural Processing Letters, 53(5), 3073-3093.

[38] Jha, S., & Mehta, A. K. (2022). A hybrid approach using the fuzzy logic system and the modified genetic algorithm for prediction of skin cancer. Neural Processing Letters, 54(2), 751-784.

[39] Araújo, R. L., de Araujo, F. H., & Silva, R. R. (2022). Automatic segmentation of melanoma skin cancer using transfer learning and fine-tuning. Multimedia Systems, 28(4), 1239-1250.

[40] Shorfuzzaman, M. (2022). An explainable stacked ensemble of deep learning models for improved melanoma skin cancer detection. Multimedia Systems, 28(4), 1309-1323.

[41] Ahmad, M., Ahmed, I., Ouameur, M. A., & Jeon, G. (2022). Classification and Detection of Cancer in Histopathologic Scans of Lymph Node Sections Using Convolutional Neural Network. Neural Processing Letters, 1-16.

[42] Kiani, K., & Sharafat, A. R. (2011). E-shaver: An improved DullRazor® for digitally removing dark and light-colored hairs in dermoscopic images. Computers in biology and medicine, 41(3), 139-145.

[43] Chao, X., & Zhang, L. (2021). Few-shot imbalanced classification based on data augmentation. Multimedia Systems, 1-9.

[44] Tschandl P., Rosendahl C. & Kittler H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Sci. Data 5, 180161 doi.10.1038/sdata.2018.161 (2018)

[45] Rotemberg, V., Kurtansky, N., Betz-Stablein, B., Caffery, L., Chousakos, E., Codella, N., Combalia, M., Dusza, S., Guitera, P., Gutman, D., Halpern, A., Helba, B., Kittler, H., Kose, K., Langer, S., Lioprys, K., Malvehy, J., Musthaq, S., Nanda, J., Reiter, O., Shih, G., Stratigos, A., Tschandl, P., Weber, J. & Soyer, P. A patient-centric dataset of images and metadata for identifying melanomas using clinical context. Sci Data 8, 34 (2021). https://doi.org/10.1038/s41597-021-00815-z