



SkinCancerNet: Automated Classification of Skin Lesion Using Deep Transfer Learning Method

Beyda Taşar 

Mechatronics Engineering Department, Firat University, Elazığ 23200, Turkey

Corresponding Author Email: btasar@firat.edu.tr

<https://doi.org/10.18280/ts.400128>

Received: 12 December 2022

Accepted: 8 February 2023

Keywords:

transfer learning, deep learning, skin cancer, skin lesions

ABSTRACT

Skin cancer has become one of the most common diseases due to the depletion of the ozone layer and the decrease in its protection. Detection and classification of skin cancer in the early stages of its development allows patients to receive appropriate treatment quickly. In this article, a modified CNN framework based on transfer learning is proposed for the classification of skin lesions from skin dermoscopy images. In the proposed framework, pre-trained CNN architectures are used. VGG16, ResNet50, DenseNet121, MobileNet, and Xception models were pre-trained using ImageNet images and training weights. In the study training and tests were performed on the HAM10000 skin lesions data set. The classification accuracy of the modified DenseNet121, VGGNet16, ResNet50, MobileNet, and Xception models were calculated as 94.29%, 93.28%, 87.10%, 83.10%, and 80.05% respectively. It was observed that the accuracy success of the proposed transfer learning framework in skin lesion type classification surpasses classical deep learning architectures.

1. INTRODUCTION

According to the studies conducted by the World Health Organization (WHO), the amount of ultraviolet radiation exposure by living things has increased due to the thinning of the ozone layer. As a result, there has been an increase in skin cancer cases [1]. Malignancy and melanoma lesions are the most common forms of skin cancer [2, 3]. These two types of lesions account for the majority of skin cancer deaths worldwide [4, 5]. Skin cancers detected at an early stage by clinical experts can be treated with surgery, radiology, and immunotherapy methods. And in the case of detection at an early stage, the survival rate of the patients is over 95%. This rate is below 15% in cases detected in advanced stages [6]. Early detection of skin cancer appears to increase the chances of survival [7, 8]. A non-invasive dermoscopy imaging method is used to improve melanoma diagnostic accuracy [9]. It has been shown that the lesion detection accuracy of experts is approximately 75% by the dermoscopy imaging method [10]. Manual interpretation of dermoscopy images is time-consuming and detection success depends on dermatologist experience and clinical education level [11]. Due to the stated problems, there is a need for computer-aided applications that can classify skin lesions with high accuracy and much faster [12-14]. In the last five years, researchers have been working on the computer-assisted diagnosis of skin cancers, especially using image processing, machine learning (ML), and deep learning (DL) techniques [15, 16]. The skin lesions studies have accelerated since opening large medical databases (ISIC 2019 and HAM10000 etc.) to researchers through competitions. Recent research reports show that computer-aided systems have a higher lesion type classification success than dermatologists, and reveal the importance of studies on this subject [17-19].

1.1 Related works

Although the use of basic machine learning algorithms was common in the first studies on the classification of skin lesions, researchers have turned to deep learning-based architectures in recent years. Studies in this field can be examined under two headings according to the data set used. While most of the researchers used widely accepted standard large-scale data sets (DERMOFIT, PH2, ISIC, HAM10000, etc.), a minority of researchers carry out skin lesion detection and classification studies on dermoscopy images that are collected individually. The use of standard datasets provides the opportunity to objectively compare the success of the methods developed by the researchers with the studies of other researchers using the same dataset. In this section, information is presented about the data set used by the researchers, the methods they developed, and the results they obtained.

A total of 5846 clinical images were collected from 3551 patients by Jinnai et al. [3]. The dataset contains images of six different skin lesions. They developed FRCNN method. The accuracy of this method was 86.2% for the six-class problem and 91.5% for the two-class (benign or malignant) problem. The classification performance of FRCNN was compared with the opinions of 20 dermatologists. As a result, FRCNN classification accuracy has been reported to be higher than that of dermatologists [3]. Kawahara et al. [20] was used pre-trained CNN architecture and achieved 85.8 percent average classification accuracy for five lesion groups in the Dermafit Image Library. They reported that the proposed method reduced the secondary training time and improved the accuracy of pre-training [20]. Liao et al. [21] used transfer learning method in VGG19 and GoogleNet architectures to increase educational success. The layers first weights were calculated with the ImageNet dataset. Using the Dermnet dataset, he determined the weights of the network again with

a secondary training. He tested the architecture he developed on the Dermnet dataset on the skin lesions dataset he received from the New York State Department of Health. It showed that the VGG19 model achieved 73.1% classification accuracy and the GoogleNet model reached 91% accuracy [21]. Pacheco et al. [22] used the ResNet50, ResNet101, GoogleNet, MobileNet, and VGG16/19 models for the classification of six different skin lesions. In their studies, they created a new dataset via a camera instead of using dermoscopy images. they reported that they achieved a classification accuracy of 78.8%, 75.7%, 77.9%, 76.2%, 74.6%, and 75.0%, respectively, with the ResNet50, ResNet101, GoogleNet, MobileNet, VGG16/19 methods [22]. Hekler et al. [23] explored the potential benefit of combining expert knowledge and AI-powered applications for skin cancer classification. They analyzed the performance of the CNN architecture, which they developed with the transfer learning method, first alone and then together with an expert dermatologist. The classification accuracy achieved by the Expert alone evaluating the skin lesion is 66%. The classification accuracy of the CNN architecture alone is 86.1%. The classification accuracy obtained by the synthesis of expert and CNN architectural results were shown to be 89% [23]. Esteva et al. [24] used a pre-trained Inception V3 model. After training on 129,450 clinical images for 2032 skin diseases, they achieved 72% classification accuracy. CNN produced performance on par with 21 skin specialists in skin lesion classification [24]. Saba et al. [1] developed a three-step methodological approach to detect skin cancer using a convolutional neural network (DCNN). In the first step, they applied HSV color conversion and contrast enhancement to the images. In the second step, lesion margin extraction was performed. In the last stage, transfer learning was carried out using the Inception V3 model. They have tested the proposed method tested on PH2, ISBI 2016, and ISBI 2017 datasets. The proposed method reached 98.4% accuracy in the PH2 dataset, 95.1% in the ISBI dataset, and 94.8% in the ISBI 2017 dataset [1]. Lopez et al. [25] adopted the transfer learning method based on the VGGNet architecture for the detection of skin cancer. Tested the success of the model on ISB 2016 skin cancer lesion images. They stated that network success reached 78.66% sensitivity [25]. Kassani et al. [26] first applied pre-processing methods to the image dataset. They improved the image quality with light correction, contrast enhancement, and decoration methods. They then performed the classification of seven types of skin lesions using the ResNet50, AlexNet, Xception, VGGNet16, and VGGNet19 architectures. They stated that ResNet50 was the most successful architecture with an accuracy rate of 92.08%. The accuracy performance of AlexNet, Xception, VGGNet16 and VGGNet19 architectures was reported as 84%, 90%, 89% and 90%, respectively [26]. Abbas et al. [27] developed a deep learning (COE-Deep) architecture to detect melanocytic and non-melanocytic (MnM) skin lesions. They used the convolutional neural network (CNN) model for feature extraction and reported an average network test success of 90% sensitivity, 93% specificity, 91.5% accuracy, and AUC=0.92 in ISIC dataset [27]. Deif et al. [28] proposed four different CNN architectures for the classification of skin lesions. These are VGG16, VGG19, MobileNet, and InceptionV3. They used the HAM10000 dataset for training, verification, and testing of the system. They reported that the accuracy rates of the VGG16, VGG19, MobileNet, and InceptionV3 architectures were 87.42%, 85.02%, 88.22%, and 89.81%, respectively [28]. Brinker et al. [29] used the

ResNet50 model for skin lesions classification problems. They trained the proposed CNN model with 12,378 dermoscopy images. They compared the performance of the CNN architecture with the results of 157 dermatologists from 12 university hospitals in Germany. The mean sensitivity and specificity of dermatologists in lesion classification were 74.1% and 60%, respectively, while the mean sensitivity and specificity of CNN were reported as 87.5% and specificity of 60% [29]. Alqudah et al. [30] used GoogleNet and AlexNet architectures to classify skin lesions into three categories (benign, melanoma, and seborrheic keratosis). The researchers used transfer learning and gradient descent adaptive momentum learning rate (ADAM) methods for training. They tested the network success using the ISIC database under two scenarios: segmented and non-segmented. They reported overall accuracy as 92.2% for the segmented dataset and 89.8% for the non-segmented dataset [30]. Thurnhofer-Hemsi and Domínguez [31] a hierarchical deep learning framework is adopted for skin cancer detection. DensNet201, GoogleNet, Inception-ResNetV2, Inception V3, and MobileNet V2 architectures were trained with the transfer learning method. HAM10000 dataset was used as a dermoscopy image data set. In addition, they used data enhancement techniques to increase performance. DensNet201, GoogleNet, Inception-ResNetV2, Inception V3, and MobileNet V2 methods success rate was expressed as 94.5%, 83.9%, 86.05%, 86.62% 88.6% and 88.34%, respectively. The results demonstrated that the DenseNet201 network is the most suitable for this task [31]. Le et al. [32] proposed an end-to-end deep learning model without preprocessing steps or feature selection. They used a modified ResNet50 deep learning model to classify skin lesion images in the HAM10000 dataset. They achieved 93% average accuracy [32]. Chaturvedi et al. [33] designed a fast-performance web application integrated with the MobileNet model for the classification skin lesions. They achieved 83.1% accuracy for HAM10000 data set [33]. MobileNet deep learning model developed by Mohamed et al. [34] for skin lesion classification reached 92.7% accuracy on the HAM10000 dataset [34]. Gupta et al. [35] used EfficientNet B1 model for the classification of skin cancer into 7 categories. They tested model performance on the HAM10000 dataset. They classified skin lesion images with a validation accuracy of 94.1%, top3 accuracy of 99.0%, and top5 accuracy of 99.9%. The weighted average of precision, recall, and f1-score of the method was found to be 0.94, 0.94, and 0.94 respectively [35]. Liu et al. [36] suggested a relation-driven semi-supervised framework for skin lesions classification. They obtained 92.54% accuracy and 60.68% F1 score with their method [36]. Al-Masni et al. [37] used an integrated diagnostic framework for the multiple (seven) class skin lesions classification. The integrated diagnostic system was tested on ISIC 2018 database. The classification performance was 88.05%, 89.28%, 87.74%, and 88.70% of by Inception-v3, ResNet-50, Inception-ResNet-v2, and DenseNet-201 models, respectively [37]. Sae-Lim et al. [38] proposed the modified MobileNet model for skin lesion classification. For the evaluation of their model, used the HAM 10000 dataset. The comparison results showed that their modified model had achieved higher accuracy, specificity, sensitivity, and F1-score than the traditional MobileNet [38]. Bassi and Gomekar [39] proposed deep-learning models for the classification of skin lesions. They used the HAM10000 dataset for testing. They reported the best accuracy of 82.8% and an average F-score of 70% [39]. Sherif et al. [40] used a CNN method for

the classification of skin lesions. The proposed models are trained and evaluated on ISIC 2018 Challenge database. They achieved an accuracy of 96.67% for the validation set [40]. Nugroho et al. [41] used a convolutional neural network (CNN) for the identification of cancer. They tested their model on HAM10000 skin cancer dataset. The training and testing accuracy of their proposed model was 80% and 78%, respectively [41]. Kassem et al. [42] proposed transfer learning method to detect skin lesions. They used pre-trained GoogleNet model. They reached 94.92% accuracy, 79.8% sensitivity, 97% specificity and 80.36% precision on the ISIC 2019 Challenge database [42].

1.2 Problem statement and contributions

Classification of multiclass skin lesions by image-based methods is a very difficult problem. The main reason for this is that different skin lesions show similarity to each other in color, the same type of lesions can occur in different forms, and the lesion appearance is variable due to differences in the skin color of individuals. In addition to these variables, surface illumination problems, veins in the skin, hair, acne, etc. The presence of tissues also affects the segmentation of the lesion. And it can reduce the accuracy of feature extraction [43]. Therefore, it is extremely important to perform effective preprocessing steps on the lesion image data. The success of the preprocessing steps significantly affects the classification success.

Although CNN architectures, which are widely used in the classification of skin lesions, achieve high performance on large data sets, the success rate decreases as the dataset size gets smaller. Therefore, by adopting the transfer learning method, the calculated weights of the pre-trained CNN architecture can be given as input to the secondary training process (training with skin lesion spatial data). This approach will be beneficial in increasing the classification success [43, 44]. Therefore, in this study, modified VGG16, ResNet50 and MobileNet, DenseNet12 and Xception network models were pre-trained with the ImageNet dataset. With the trained network model framework, the HAM10000 skin lesion visual dataset was classified faster and with higher accuracy.

The main contribution and prominent aspects of this study are as follows:

- A deep learning framework based on the transfer learning principle was established for the detection/classification of seven types of skin lesions. Lesions to be detected are Melanocytic nevus (nv), Melanoma (mel), Basal cell carcinoma (bcc), Actinic keratosis (akiec), Vascular lesions (vasc), Benign keratosis-like lesions (bkl), Dermatofibroma (df).
- Five different architectures modified architectures VGG16, MobileNet, ResNet50, DenseNet121, and Xception were used for transfer learning.
- Gaussian noise is applied to the input to mitigate overfitting.
- The batch normalization method was used in the proposed diagnostic frame. In this way, significantly reduces the number of training cycles required to build standardized and deep networks.
- Six different performance evaluation metrics were calculated and the success of the proposed diagnostic framework was tested and compared.

The organization of current study is three main parts. In Section 2 the proposed method, information on transfer learning, modified network architectures and feature layers, and performance metrics are included. In Section 3 classification results for the five proposed methods, confusion matrix, and performance table are presented. In addition, the results of previous studies and the performance of the proposed method are compared in this section. In the last section, all the important results of the study are summarized with data.

2. MATERIAL AND METHOD

This section introduces the proposed transfer learning framework for the classification of skin lesions in dermoscopy images. The five well-known CNN architectures: VGG16, ResNet50, DenseNet121, and Xception were pre-trained using the ImageNet dataset in the first step, then fine-tuned with dermoscopy images. The proposed transfer learning framework is shown in the block diagram of Figure 1.

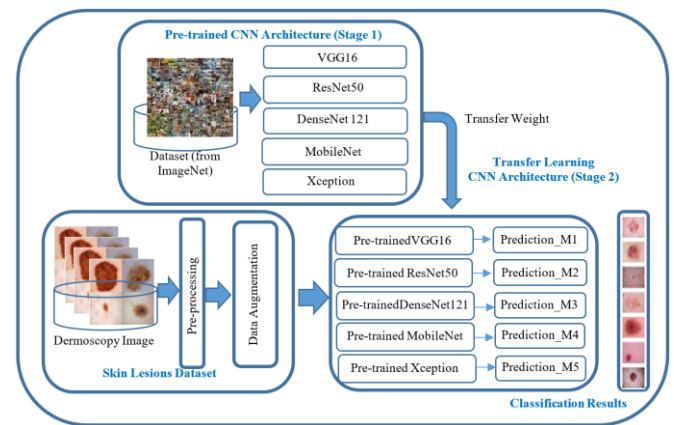


Figure 1. Proposed transfer learning framework for skin lesion classification

2.1 Data pre-processing and augmentation

In this study, the Skin Cancer MNIST: HAM10000 dataset, which was presented to the researchers as part of the competition on the Kaggle platform, was used [45]. In the HAM10000 metadata format, each of the images is mainly sized at (600×450×3) [44]. This study collected lesions from seven categories in two groups: Benign: Melanocytic Nevi (nv), Vascular Lesions (vasc), Dermatofibroma (df), and benign keratosis (bkl) Malignant: melanoma(mel), basal cell carcinoma (bcc), actinic keratoses (akiec) (Figure 2).

In order to make all images as the same size, and remove various types of noise, the preprocessing phase is important. For this reason, to lessen the computational cost of our proposed architecture, we resized and rescaled the image size from 224×224×3. To achieve better efficiency and accuracy with CNN requires large datasets. In addition, CNN will smooth the output with tiny datasets as overfitting will occur. But also, overfitting means that the network performs very well on training data, but performs poorly on the test data. Due to the fact that dermoscopy images are rotationally stable, images of skin lesions can easily be analyzed from various angles without any diagnostic changes [46].

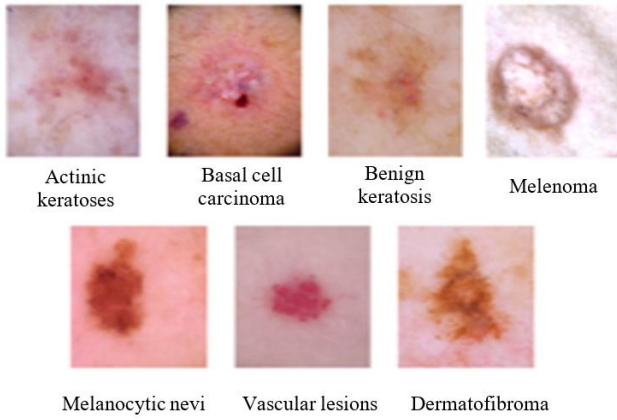


Figure 2. Skin lesions

The data augmentation techniques were implemented in the proposed system to increase the dataset and minimize overfitting issues. Some geometric transformation methods were used for datasets and the number of samples was increased using basic image processing techniques for data magnification. The main methods used to reproduce data are color processing, transform (flip, scale, and rotate), translation and noise perturbation, etc. The methods and parameter values used in data augmentation are presented in Table 1 in detail.

Table 1. Hyper parameters used in data augmentation

Hyper parameter	Value
rotation range	45 degree
Width shift range	0.1
Height shift range	0.1
Shear range	0.01
Zoom range	[0.9, 1.25]
Brightness range	[0.7, 1.3]
Horizontal flip	True
Fill mode	'reflect'
Data format	'channels_last'

2.2 Transfer learning

Large quantities of information are required for the training of a CNN from scratch, but in some situations, a large dataset is very difficult to organize. Contrary to an ideal scenario, this is not the case for most actual applications.

Additionally, it is a complicated challenge to obtain matching training and test data. This led to the development of the idea of transfer learning. In its respective datasets, the base network is initially trained for a given job and then transferred to the target role established by the target dataset [47]. Two key steps can be used to select the pre-trained model, the size of the issue and the similarity. The pre-trained model is chosen based on a related problem that is applicable to the objective. If the target data set is smaller (i.e., less than 10,000 images) than the source database set, then the probability of overfitting is high. Likewise, if the target data is larger and corresponds to the source data sets, there would be a small overfit risk and only the previously trained model needs to be refined.

Table 2 summarizes the proposed framework for transfer learning. As a transfer learning model, five different CNN models were used separately. Gaussian noise added to input to mitigate overfitting. As it is a regularization layer, it is only active at training time. In the transfer learning base model, the convolution layers remained constant and their weights were transferred. The batch normalization layer was used after the

transfer learning model. Normalization batch technology is a very deep neural network training technique that standardizes the inputs for each mini-batch in a layer. This stabilizes the learning process and significantly decreases the number of training cycles required to form deep networks [48]. Subsequently, two consecutive dropouts and dense layers were generated by a new transfer learning system.

CNN is a special form of neural network (NN) which is designed to learn images' visual features. Currently, the deeper learning approach is the most effective for image classification [49].

Table 2. Proposed transfer learning frame

Layer Name	Layer Type
Input Layer	Image_RGB (224x224x3)
Noise Layer	gaussian_noise=0.05
Based Pre-Trained Model	Dense121, VGG16, MobileNet, Reenet50, Xception
Normalization method	batch_normalization
Dropout Layer 1	layers.Dropout(0.5)(base_layer)
Flatten Layer 1	layers.Flatten()(dropout_layer_1)
Dense Layer 1	layers.Dense(256, activation="relu")(flat_layer)
Dropout Layer 2	layers.Dropout(0.5)(dense_1)
Dense Layer 2	layers.Dense(256, activation="relu")(dropout_layer_2)
Output Layer	outputs = layers.Dense(7, activation="sigmoid")(dense_2)

While numerous CNN architectures have been used to detect skin cancer and benign of skin lesions, large quantities of data are difficult to obtain for the training of a CNN. Transfer learning [34] is known to solve this problem with the partial reuse of a model which has been trained for a special resource task. Separate CNN architectures are originally used to isolate the functionality and for classification tasks, they are merged into a fully-linked layer. Multiple properties derived from a single identifier can be found in combined properties such as circularity, roundness, compactness, etc., that the form descriptor can reflect. The five most known and newest CNN networks are DenseNet 121 [31], VGGNet16 [50] and ResNet50 [51], MobileNet [52], and Xception [53], for the classification of skin lesions. The following architectures have been prepared for different general image descriptors, followed by a feature removal using transfer learning theory from dermoscopy images. Below is an overview of the basic structure of each used CNN architecture.

VGGNet: VGGNet parallels AlexNet with the exception of additional layers of convolution. VGGNet contains 13 convolutions, smoothing, pooling, and three fully connected layers, total of 16 layers, which are completely related. The mesh uses a filter scale of 3 to 3 windows and a 2 to 2 mesh. It also involves batch normalization, non-linear ReLU activations, and layer pooling after two or three turns. It extracts 25,088 image characteristics for the classifier.

DenseNet 121: DenseNet falls under the classical network group. By using a composite function operation, an output of the previous layer is an entry of the second layer. It consists of a convolution layer, a batch-standardization layer, and a non-linear activation layer. These ties make the network directly linked with $L(L+1)/2$. L is the architectural number of layers. There are many variants of DenseNet, such as DenseNet-121, DenseNet-160, DenseNet-201, etc. The numbers reflect the number of neural network layers. DenseNet121 consists of

five layers of invitation and pooling, and three layers of the transaction (6,12,24) (1x1 and 3x3 conv).

Resnet: ResNet is a deep residual network which, in the ImageNet classification mission, achieves good results. Thanks to its deep structures, ResNet integrates many convolution filters that handle degradation problems and reduce the time required for training. ResNet-50/101: the residual layer functions in respect of layer inputs are reformulated by this model instead of using unreferenced functions learning. It is achieved by means of an approach that skips the relationships of non-linearity parts of certain layers and standards (ReLU). In this work, we use two ResNet versions, one containing 49 convolution layers and the other 100. The classifier is returned by both extractors with 2,048 image features.

MobileNet: this CNN architecture consists of deeply separable convolutions, consisting of a profound confusion and point-specific convolution. Compared to networks using regular convolutions of the same depth, this technique greatly decreases the number of parameters. Two hyper parameters are also included in the model to monitor the scale. In this job, we use the complete MobileNet, leading to 22 layers of convolution. It also uses ReLU and batch standardization, much like the others.

Xception: The Xception deep neural network which stands for extreme inception, was made by François Chollet [53]. Xception architecture has depth-wise separable convolutions. Xception has 36 convolutional layers to extract important features and is inspired by Inception Chollet wherein the Inception modules are replaced with depth-wise separable convolutions consisting of a depth-wise convolution [53].

2.3 Model hyperparameters optimization

Pre-trained model's hyperparameters were tried to be optimized using the grid search method to obtain the highest performance. Batch sizes in the range of 10-100 and epoch values in the range of 50-150 were examined to find the batch

size and epoch value that would provide the best performance. In addition, the effects on performance were analyzed in seven different optimizers. In addition, the effect of a learning rate of [0.0001: 0.3], and a change in momentum values of [0.0:0.9] on network performance was examined. batch size 64, 100 epochs, Adam optimizers, 0.0001 learning rate and 0.00 momentum parameters, which were determined to give the most optimum results, were selected to be used in the study.

2.4 Performance evaluation methods

In cases where there is an imbalance between the classes in the data set, knowing the accuracy value alone does not constitute an objective approach in the evaluation of performance. For this reason, seven different metric values were calculated in this study. These criteria are precision, sensitivity, accuracy, specificity, F1-score, Matthews correlation coefficient, and Kappa [54]. Table 3 contains the formulas and brief explanations of these criteria. The TP in the table represents the true positive, that is, the number of correctly predicted samples. FN is the number of incorrectly guessed samples. TN (true negative) is the number of correctly predicted negative samples and FP (false positive) is the number of incorrectly predicted negative samples. Especially in the medical area, sensitivity, precision, and specificity values are important. Sensitivity refers to the probability of a positive test. Specificity is the calculation of successfully labeled non-skin lesion marks. Precision or a positive predictive value tests the correct percentage of classified marks. The number of skin lesions divided by the total number of skin lesions is shown with the proper classification accuracy (ACC). In addition, the confusion matrix for all models was calculated. A confusion matrix is a numerical table used to demonstrate the classification model output effects on the test data known from the goal labels. As a visual table, we use the confusion matrix to show how our methods predict our data.

Table 3. Lookup table of performance evaluation metrics used in this study

Performance Metric	Acronym	Equation	Explanation
Positive Predictive Value	PPV Precision	$\frac{TP}{TP + FP}$	The ratio of positive samples that are predicted correctly out of all the samples predicted to be positive.
Negative Predictive Value	NPV	$\frac{TN}{TN + FN}$	The ratio of negative samples that are predicted correctly out of all the samples that are predicted to be negative.
True Positive Rate	TPR Sensitivity	$\frac{TP}{TP + FN}$	The ratio of TP outcomes to the total number of actual positive samples.
True Negative Rate	TNR Specificity	$\frac{TN}{TN + FP}$	The ratio of TP outcomes to the total number of actual negative samples.
Accuracy	ACC	$\frac{TP + TN}{TP + FN + TN + FP}$	The ratio of TP outcomes to the total number of actual positive samples.
Multi class Accuracy	ACC	$\frac{TP + TN + FP + FN}{TP + TN + FP + FN}$	The ratio of the number of correct predictions made by the method out of the total number of predictions made.
F1-Score	F1	$2 \times \frac{PPV * TPR}{PPV + TPR}$	The weighted average between the PPV and TPR scores.
MCC-Matthews correlation coefficient	MCC	$\max \left(\frac{(TP * TN) - (FP * FN)}{((TP + FP) * P * N * (TN + FN))^{0.5}}, \frac{((TP + FP) * P * N * (TN + FN))^{0.5}}{(TP * TN) - (FP * FN)} \right)$	
Cohen's kappa	Kappa	$pe = \frac{(P * (TP + FP) + (FN + TN))}{(TP + TN + FP + FN)^2}$ $kappa = \max \left(\left \frac{po - pe}{1 - pe} \right , \left \frac{pe - po}{1 - po} \right \right)$	
Receiver Operating Characteristic	ROC	The probability curve	A line from (0.0) to (1.1) in coordinates of TPR and FPR. which shows TPR against FPR at different thresholds.
The value of Area Under the ROC Curve	AUC	$\int_0^1 ROC(t) dt$	The numerical index represents the area under the ROC curve.

3. EXPERIMENTAL RESULTS AND DISCUSSION

3.1 Data set and exploratory data analysis

The number of samples present for every type of skin lesion in the augmented dataset is presented in Table 4. For skin lesion class 80% of the images are used for training and 20% are used for test.

Table 4. The number of samples for each type of skin lesion

Class Name	Clinical diagnosis	Total	Training	Test
Class 1	Melanocytic nevi (nv)	5030	4024	1006
Class 2	Melanoma (mel)	880	704	176
Class 3	Basal cell carcinoma (bcc)	390	312	78
Class 4	Actinic keratoses (akiec)	235	188	47
Class 5	Vascular lesions (vasc)	110	88	22
Class 6	Benign keratosis-like lesions (bkl)	800	640	160
Class 7	Dermatofibroma (df)	75	60	15
Total		7.520	5.985	1.504

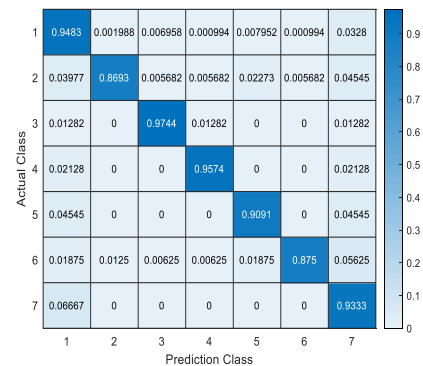
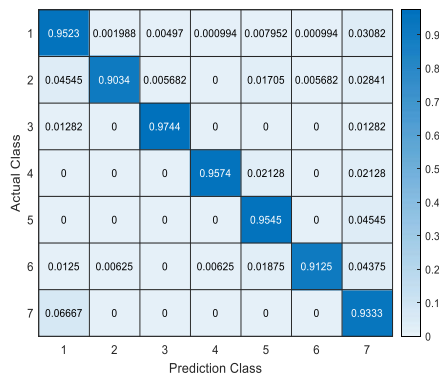
This dataset train-test split was made using tf.keras randomly split function. The study was carried out on a computer with Intel (R) Core (TM) i7-7700HQ CPU 16 GB RAM memory, operating speed of 2.81 Ghz. In addition,

external GPU support provided by the Kaggle platform is used. In this study, all process was coded in the python programming language. And it used high-level tf.keras API and Tensor Flow 2.3.0 framework.

3.2 Results and discussion

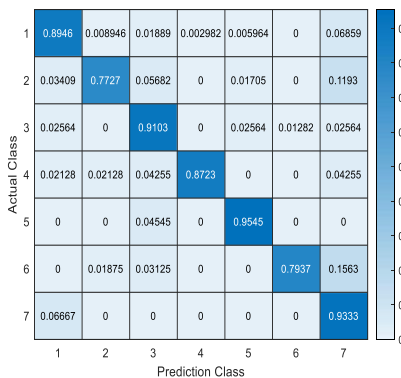
Within the scope of the study, the transfer learning method was applied to the five most popular CNN models for the detection of skin lesions. The performance of the models was mutually analyzed and the results are presented in detail in this section. There is a confusion matrix for each model in Figure 3 and a ROC curve in Figure 4. The AUC - ROC curve is a performance measurement for the classification problem at different threshold values.

Table 5 shows ACC, sensitivity, specificity, precision, F-score, kappa, and Matthews' correlation coefficient values, which show the classification success of each model. For each model, the highest classification success was obtained for the class 3, 5, and 7 lesions category, while the lowest success was obtained in the classification of lesions belonging to the second and sixth classes. The average classification accuracy for all lesion classes via the VGGNet, ResNet50, DenseNet121, MobileNet, and Xception architectures was 94.29%, 93.28%, 87.10%, 83.10%, 80.05% respectively. While DenseNet121 has the highest success, the Xception model has the lowest classification performance.

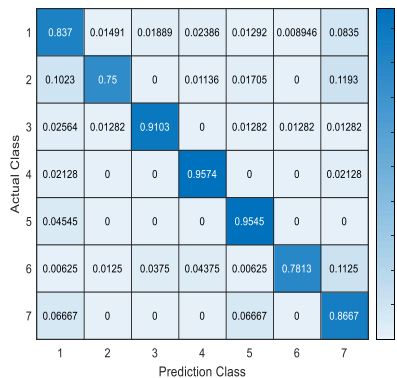


(A) DenseNet121

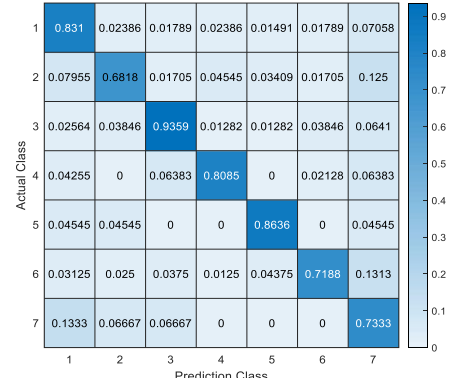
(B) VGG16



(C) ResNet50



(D) MobileNet



(E) Xception

Figure 3. Confusion matrices for each architecture

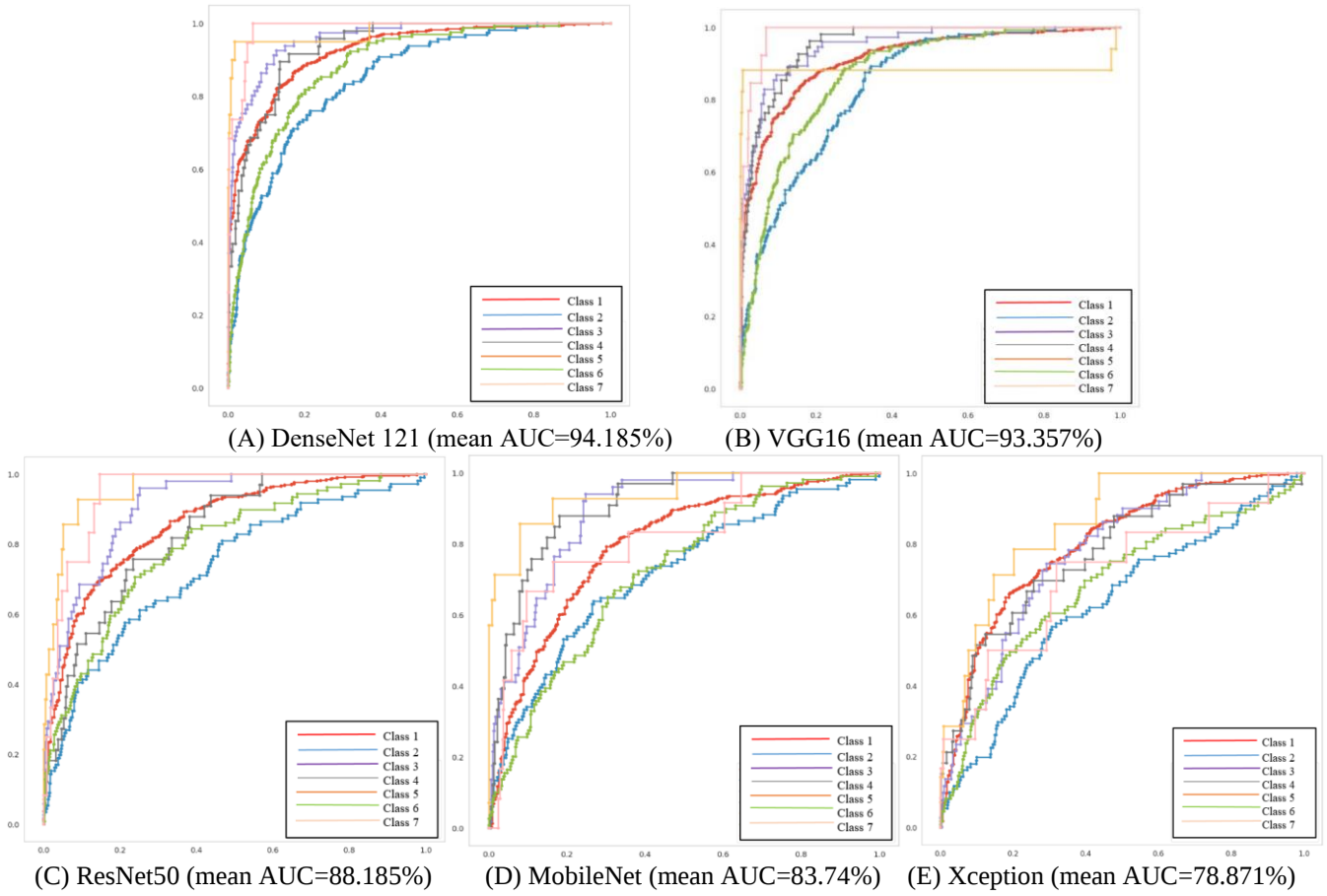


Figure 4. ROC curves for each architecture

Table 5. Performance measurements for each architecture

	Class No	Accuracy of Single	Specificity	Precision	F-Score	Kappa	Matthews Correlation Coefficient
DenseNet121	Class 1 (nv)	0.9522	0.9771	0.8738	0.9113	0.7061	0.8970
	Class 2 (mel)	0.8983	0.9986	0.9909	0.9423	0.7286	0.9347
	Class 3 (bcc)	0.9743	0.9982	0.9891	0.9817	0.7175	0.9787
	Class 4 (akiec)	0.9574	0.9987	0.9924	0.9746	0.7206	0.9707
	Class 5 (vasc)	0.9545	0.9891	0.9362	0.9452	0.7143	0.9361
	Class 6 (bkl)	0.9125	0.9988	0.9927	0.9509	0.7280	0.9442
	Class 7 (df)	0.9333	0.9696	0.8364	0.8822	0.7040	0.8631
	Multiclass	0.9429	0.9901	0.9446	0.9412	0.7665	0.9321
VGGNet	Class 1 (nv)	0.94831	0.97384	0.98656	0.96706	0.19089	0.90570
	Class 2 (mel)	0.87429	0.99699	0.97452	0.92169	0.78116	0.91369
	Class 3 (bcc)	0.96203	0.99368	0.89412	0.92683	0.89134	0.92330
	Class 4 (akiec)	0.95745	0.99725	0.91837	0.93750	0.93626	0.93566
	Class 5 (vasc)	0.90909	0.98987	0.57143	0.70175	0.96226	0.71587
	Class 6 (bkl)	0.88050	0.99851	0.98592	0.93023	0.80121	0.92432
	Class 7 (df)	0.9333	0.96438	0.22059	0.36145	0.94513	0.46123
	Multiclass	0.9328	0.9875	0.7913	0.8180	0.7256	0.8213
ResNet50	Class 1 (nv)	0.89463	0.97992	0.98901	0.93946	0.13545	0.84195
	Class 2 (mel)	0.77273	0.99021	0.91275	0.83692	0.78792	0.82088
	Class 3 (bcc)	0.91026	0.97405	0.65741	0.76344	0.87802	0.75955
	Class 4 (akiec)	0.87234	0.99794	0.93182	0.90110	0.93968	0.89854
	Class 5 (vasc)	0.95455	0.99258	0.65625	0.77778	0.96422	0.78798
	Class 6 (bkl)	0.79375	0.99926	0.99219	0.88194	0.81063	0.87624
	Class 7 (df)	0.93333	0.92008	0.10526	0.18919	0.90245	0.29868
	Multiclass	0.8710	0.9791	0.7492	0.7557	0.4733	0.7548
MobileNet	Class 1 (nv)	0.83698	0.95171	0.97229	0.89957	0.07257	0.75085
	Class 2 (mel)	0.75000	0.98644	0.88000	0.80982	0.78784	0.79001
	Class 3 (bcc)	0.92208	0.98247	0.73958	0.82081	0.88606	0.81556
	Class 4 (akiec)	0.95745	0.97734	0.57692	0.72000	0.91761	0.73347
	Class 5 (vasc)	0.95455	0.98717	0.52500	0.67742	0.95895	0.70268
	Class 6 (bkl)	0.78125	0.99255	0.92593	0.84746	0.80678	0.83468
	Class 7 (df)	0.86667	0.91599	0.09420	0.16993	0.89916	0.26941
	Multiclass	0.8310	0.9705	0.6734	0.7064	0.3099	0.6995

Xception	Class 1 (nv)	0.83101	0.94882	0.96984	0.89507	0.06154	0.74363
	Class 2 (mel)	0.68182	0.97534	0.78431	0.72948	0.78951	0.69884
	Class 3 (bcc)	0.82955	0.97826	0.70192	0.76042	0.87515	0.74728
	Class 4 (akiec)	0.80851	0.97614	0.52055	0.63333	0.92177	0.63525
	Class 5 (vasc)	0.86364	0.98056	0.39583	0.54286	0.95411	0.57657
	Class 6 (bkl)	0.71875	0.98154	0.82143	0.76667	0.80671	0.74317
	Class 7 (df)	0.73333	0.91795	0.08209	0.14765	0.90263	0.22710
	Multiclass	0.8005	0.9655	0.6109	0.6394	0.1855	0.6245

Table 6. Comparison of performance for skin lesion detection

	Method	Accuracy %	Sensitivity %	Specificity %
(A) Published studies used HAM10000 dataset				
[28]	Dilated InceptionV3	89	89	89
[29]	CNN	-	84.5	86.5
[31]	DenseNet201	94.5	-	-
[32]	Ensemble Model	93.00	86.00	82.00
[33]	MobileNet	92.70	87.00	81.00
[34]	MobileNet	92.70	87	81
[35]	EfficientNetB1	94.00	94	94
[39]	VGG16	82.80		64.57
[41]	From scratch	80.93	68.97	53.95
(B) Skin Cancer Kaggle Challenge results used HAM 100000 Collected Dataset [45, 55] (not published as paper)				
Nils et al.	Ensemble of multi-res EfficientNets with SEN154 2	92.6	50.7	97.7
Zhou et al.	Ensemble of EfficientNetB3-B4-Seresnext101	91.7	60.7	95.2
Pacheco et al.	Ensemble classifiers	91.9	50.7	96.5
Chouhan	Densenet-121	91.0	47.3	96.7
Dat et al.	CNNs based on Inception-ResNet. XceptionNet And EfficientNet	91.4	55.5	95.0
Zhang	Malanet based on DenseNet	89.7	66.6	91.6
Xing et al.	Class-centroid-based openset ensemble	91.9	55.7	95.1
Subhranil et al.	Long-tail distribution based classifiers	91.3	49.7	95.8
Zadeh et al.	Softmax ensemble and sigmoid ensemble classifier model	92.0	51.9	95.6
Cohen et al.	Test time augmentation on ensemble models	92.4	46.9	96.3
Sara et al.	Xception. Inception-ResNet-V2. NasNetLarge	92.1	46.0	96.2
	DenseNet121	94.29	94.04	99.01
	VGG16	93.28	92.36	98.75
Our Study	Resnet50	87.10	85.79	97.91
	MobileNet	83.10	86.70	97.05
	Xception	80.05	78.09	96.55

In Table 6 similar studies on the classification of skin lesions and the results obtained from the current study are compared in terms of accuracy, sensitivity, and specificity criteria. Table 6 is divided into two main sections. The first contains the published papers results of the studies using HAM10000 dataset. The second part includes the results of the researchers who participated in the "Skin Lesion Analysis Towards Melanoma Detection" competition held on the Kaggle platform and showed the best performance using the shared dataset. In this study, with DenseNet 121 and VGG16 network architectures whose transfer learning basis was changed, accuracy levels of 94.29% and 93.28% were achieved, respectively. When the results of the study are compared with studies using the HAM10000 dataset directly, it is seen that the obtained accuracy rate is only 0.21% lower than the 94.5% accuracy rate obtained in [31] but considerably higher than other studies. Sensitivity and specificity values by DenseNet121 were obtained as 94.04% and 99.01% respectively, which is the highest value compared to the studies in the literature.

4. CONCLUSIONS

This paper uses the principle of transfer learning. We proposed a novel deep-learning framework for the classification of skin lesions. The principle of data augmentation was also suggested to maximize the performance of the CNN structure. Finally, the efficiency of the proposed system was contrasted with other current approaches and findings from the literature. Accuracy for DenseNet 121, VGGNet, ResNet50, MobileNet, and Xception was 94.29%, 93.28%, 87.10%, 83.10%, 80.05% respectively. The best classification results were obtained from DenseNet121 and the Xception model had the lowest classification results of all models. The proposed transfer learning system has been shown to provide excellent results in terms of accuracy without scratch training, which increases classification performance. It is thought that the proposed model approach will help physicians in the rapid and accurate detection of skin cancer lesions.

ACKNOWLEDGMENT

This study was supported by Firat University within the scope of MF 21.14 Graduate BAP Project.

REFERENCES

- [1] Saba, T., Khan, M.A., Rehman, A., Marie-Sainte, S.L. (2019). Region extraction and classification of skin cancer: A heterogeneous framework of deep CNN features fusion and reduction. *Journal of Medical Systems*, 43(9): 289. <https://doi.org/10.1007/s10916-019-1413-3>
- [2] Schadendorf, D., van Akkooi, A.C., Berking, C., Griewank, K.G., Gutzmer, R., Hauschild, A., Ugurel, S. (2018). Melanoma. *The Lancet*, 392(10151): 971-984. [https://doi.org/10.1016/S0140-6736\(18\)31559-9](https://doi.org/10.1016/S0140-6736(18)31559-9)
- [3] Jinnai, S., Yamazaki, N., Hirano, Y., Sugawara, Y., Ohe, Y., Hamamoto, R. (2020). The development of a skin cancer classification system for pigmented skin lesions using deep learning. *Biomolecules*, 10(8): 1123. <https://doi.org/10.3390/biom10081123>
- [4] Rogers, H.W., Weinstock, M.A., Feldman, S.R., Coldiron, B.M. (2015). Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the US population, 2012. *JAMA Dermatology*, 151(10): 1081-1086. <https://doi.org/10.1001/jamadermatol.2015.1187>
- [5] Trovitch, P., Gupta, A., Ciftci, K. (2002). Early detection and treatment of skin cancer. *Turkish Journal of Cancer*, 32(4): 129.
- [6] Balch, C.M., Gershenwald, J.E., Soong, S.J., et al. (2009). Final version of 2009 AJCC melanoma staging and classification. *Journal of Clinical Oncology*, 27(36): 6199. <https://doi.org/10.1200/JCO.2009.23.4799>
- [7] Khan, M.A., Sharif, M., Akram, T., Damaševičius, R., Maskeliūnas, R. (2021). Skin lesion segmentation and multiclass classification using deep learning features and improved moth flame optimization. *Diagnostics*, 11(5): 811. <https://doi.org/10.3390/diagnostics11050811>
- [8] Nasir, M., Attique Khan, M., Sharif, M., Lali, I.U., Saba, T., Iqbal, T. (2018). An improved strategy for skin lesion detection and classification using uniform segmentation and feature selection based approach. *Microscopy Research and Technique*, 81(6): 528-543. <https://doi.org/10.1002/jemt.23009>
- [9] Binder, M., Schwarz, M., Winkler, A., Steiner, A., Kaidir, A., Wolff, K., Pehamberger, H. (1995). Epiluminescence microscopy: A useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists. *Archives of Dermatology*, 131(3): 286-291. <https://doi.org/10.1001/archderm.131.3.286>
- [10] Abbas, Q., Garcia, I.F., Rashid, M. (2010). Automatic skin tumour border detection for digital dermoscopy using a new digital image analysis scheme. *British Journal of Biomedical Science*, 67(4): 177-183. <https://doi.org/10.1080/09674845.2010.11730316>
- [11] Yu, L., Chen, H., Dou, Q., Qin, J., Heng, P.A. (2016). Automated melanoma recognition in dermoscopy images via very deep residual networks. *IEEE Transactions on Medical Imaging*, 36(4): 994-1004. <https://doi.org/10.1109/TMI.2016.2642839>
- [12] Saba, T. (2020). Recent advancement in cancer detection using machine learning: Systematic survey of decades, comparisons and challenges. *Journal of Infection and Public Health*, 13(9): 1274-1289. <https://doi.org/10.1016/j.jiph.2020.06.033>
- [13] Barata, C., Celebi, M.E., Marques, J.S. (2018). A survey of feature extraction in dermoscopy image analysis of skin cancer. *IEEE Journal of Biomedical and Health Informatics*, 23(3): 1096-1109. <https://doi.org/10.1109/JBHI.2018.2845939>
- [14] Rehman, A., Khan, M.A., Mehmood, Z., Saba, T., Sardaraz, M., Rashid, M. (2020). Microscopic melanoma detection and classification: A framework of pixel-based fusion and multilevel features reduction. *Microscopy Research and Technique*, 83(4): 410-423. <https://doi.org/10.1002/jemt.23429>
- [15] Nasr-Esfahani, E., Samavi, S., Karimi, N., Soroushmehr, S.M.R., Jafari, M.H., Ward, K., Najarian, K. (2016). Melanoma detection by analysis of clinical images using convolutional neural network. In 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 1373-1376. <https://doi.org/10.1109/EMBC.2016.7590963>
- [16] Korotkov, K., Garcia, R. (2012). Computerized analysis of pigmented skin lesions: A review. *Artificial Intelligence in Medicine*, 56(2): 69-90. <https://doi.org/10.1016/j.artmed.2012.08.002>
- [17] Brinker, T.J., Hekler, A., Enk, A.H., Klode, J., Hauschild, A., Berking, C., Schilling, B., Haferkamp, S., Schadendorf, D., Fröhling, S., Utikal, J.S., Schrüfer, P. (2019). A convolutional neural network trained with dermoscopic images performed on par with 145 dermatologists in a clinical melanoma image classification task. *European Journal of Cancer*, 111: 148-154. <https://doi.org/10.1016/j.ejca.2019.02.005>
- [18] Fujisawa, Y., Otomo, Y., Ogata, Y., Nakamura, Y., Fujita, R., Ishitsuka, Y., Watanabe, R., Okiyama, N., Fujimoto, M. (2019). Deep-learning-based, computer-aided classifier developed with a small dataset of clinical images surpasses board-certified dermatologists in skin tumour diagnosis. *British Journal of Dermatology*, 180(2): 373-381. <https://doi.org/10.1111/bjd.16924>
- [19] Khan, M.A., Akram, T., Sharif, M., Saba, T., Javed, K., Lali, I.U., Tanik, U.J., Rehman, A. (2019). Construction of saliency map and hybrid set of features for efficient segmentation and classification of skin lesion. *Microscopy Research and Technique*, 82(6): 741-763. <https://doi.org/10.1002/jemt.23220>
- [20] Kawahara, J., BenTaieb, A., Hamarneh, G. (2016). Deep features to classify skin lesions. In 2016 IEEE 13th international symposium on biomedical imaging (ISBI), pp. 1397-1400. <https://doi.org/10.1109/ISBI.2016.7493528>
- [21] Liao, H. (2015). A deep learning approach to universal skin disease classification. *Computer Science*, pp. 1-4.
- [22] Pacheco, A.G., Krohling, R.A. (2020). The impact of patient clinical information on automated skin cancer detection. *Computers in Biology and Medicine*, 116: 103545. <https://doi.org/10.1016/j.compbiomed.2019.103545>
- [23] Hekler, A., Utikal, J.S., Enk, A.H., et al. (2019). Superior skin cancer classification by the combination of human and artificial intelligence. *European Journal of Cancer*, 120: 114-121. <https://doi.org/10.1016/j.ejca.2019.07.019>
- [24] Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter,

- S.M., Blau, H.M., Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639): 115-118. <https://doi.org/10.1038/nature21056>
- [25] Lopez, A.R., Giro-i-Nieto, X., Burdick, J., Marques, O. (2017). Skin lesion classification from dermoscopic images using deep learning techniques. In 2017 13th IASTED International Conference on Biomedical Engineering (BioMed), pp. 49-54. <https://doi.org/10.2316/P.2017.852-053>
- [26] Kassani, S.H., Kassani, P.H. (2019). A comparative study of deep learning architectures on melanoma detection. *Tissue and Cell*, 58: 76-83. <https://doi.org/10.1016/j.tice.2019.04.009>
- [27] Abbas, Q. (2017). Development of a clinically-oriented expert system for differentiating melanocytic from non-melanocytic skin lesions. *International Journal of Advanced Computer Science and Applications*, 8(7): 24-29. <https://doi.org/10.14569/ijacsa.2017.080704>
- [28] Deif, M.A., Hammam, R.E. (2020). Skin lesions classification based on deep learning approach. *Journal of Clinical Engineering*, 45(3): 155-161. <https://doi.org/10.1097/jce.0000000000000405>
- [29] Brinker, T.J., Hekler, A., Enk, A.H., Klode, J., Hauschild, A., Berking, C., Schilling, B., Haferkamp, S., Schadendorf, D., Holland-Letz, T., Utikal, J.S., Schröder, P. (2019). Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task. *European Journal of Cancer*, 113: 47-54. <https://doi.org/10.1016/j.ejca.2019.04.001>
- [30] Alqudah, A.M., Alquraan, H., Qasmieh, I.A. (2019). Segmented and non-segmented skin lesions classification using transfer learning and adaptive moment learning rate technique using pretrained convolutional neural network. *Journal of Biomimetics, Biomaterials and Biomedical Engineering*, 42: 67-78. <https://doi.org/10.4028/www.scientific.net/JBBBE.42.67>
- [31] Thurnhofer-Hemsi, K., Domínguez, E. (2021). A convolutional neural network framework for accurate skin cancer detection. *Neural Processing Letters*, 53(5): 3073-3093. <https://doi.org/10.1007/s11063-020-10364-y>
- [32] Le, D.N., Le, H.X., Ngo, L.T., Ngo, H.T. (2020). Transfer learning with class-weighted and focal loss function for automatic skin cancer classification. *arXiv preprint* arXiv:2009.05977. <http://arxiv.org/abs/2009.05977>
- [33] Chaturvedi, S.S., Gupta, K., Prasad, P.S. (2021). Skin lesion analyser: An efficient seven-way multi-class skin cancer classification using MobileNet. In *Advanced Machine Learning Technologies and Applications: Proceedings of AMLTA 2020*, pp. 165-176. https://doi.org/10.1007/978-981-15-3383-9_15
- [34] Mohamed, E.H., El-Behaidy, W.H. (2019). Enhanced skin lesions classification using deep convolutional networks. In 2019 Ninth International Conference on Intelligent Computing and Information Systems (ICICIS), pp. 180-188. <https://doi.org/10.1109/ICICIS46948.2019.9014823>
- [35] Gupta, H., Bhatia, H., Giri, D., Saxena, R., Singh, R. (2020). Comparison and analysis of skin lesion on pretrained architectures. *International Research Journal of Engineering and Technology (IRJET)*, 7(7): 2704-2707. <https://doi.org/10.13140/RG.2.2.32161.43367>
- [36] Liu, Q., Yu, L., Luo, L., Dou, Q., Heng, P.A. (2020). Semi-supervised medical image classification with relation-driven self-ensembling model. *IEEE Transactions on Medical Imaging*, 39(11): 3429-3440. <https://doi.org/10.1109/TMI.2020.2995518>
- [37] 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. <https://doi.org/10.1016/j.cmpb.2020.105351>
- [38] Sae-Lim, W., Wettayaprasit, W., Aiyarak, P. (2019). Convolutional neural networks using MobileNet for skin lesion classification. In 2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE), pp. 242-247. <https://doi.org/10.1109/JCSSE.2019.8864155>
- [39] Bassi, S., Gomekar, A. (2019). Deep learning diagnosis of pigmented skin lesions. In 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp. 1-6. <https://doi.org/10.1109/ICCCNT45670.2019.8944601>
- [40] Sherif, F., Mohamed, W.A., Mohra, A.S. (2019). Skin lesion analysis toward melanoma detection using deep learning techniques. *International Journal of Electronics and Telecommunications*, 65(4): 597-602. <https://doi.org/10.24425/ijet.2019.129818>
- [41] Nugroho, A.A., Slamet, I., Sugiyanto. (2019). Skin cancer identification system of HAM10000 skin cancer dataset using convolutional neural network. In *AIP Conference Proceedings*, 2202(1): 020039. <https://doi.org/10.1063/1.5141652>
- [42] 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. <https://doi.org/10.1109/ACCESS.2020.3003890>
- [43] Yao, Q., Guan, Z., Zhou, Y., Tang, J., Hu, Y., Yang, B. (2009). Application of support vector machine for detecting rice diseases using shape and color texture features. In 2009 International Conference on Engineering Computation, pp. 79-83. <https://doi.org/10.1109/ICEC.2009.73>
- [44] Sadeghi, M., Lee, T.K., McLean, D., Lui, H., Atkins, M.S. (2013). Detection and analysis of irregular streaks in dermoscopic images of skin lesions. *IEEE Transactions on Medical Imaging*, 32(5): 849-861. <https://doi.org/10.1109/TMI.2013.2239307>
- [45] Competition. (2019). In SI IM-ISIC melanoma classification identify melanoma in lesion images. <https://www.kaggle.com/c/siim-isic-melanoma-classi>
- [46] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y., Berg, A.C. (2016). Ssd: Single shot multibox detector. In *Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part I 14*, pp. 21-37. https://doi.org/10.1007/978-3-319-46448-0_2
- [47] Yang, L., Hanneke, S., Carbonell, J. (2013). A theory of transfer learning with applications to active learning. *Machine Learning*, 90: 161-189. <https://doi.org/10.1007/s10994-012-5310-y>
- [48] Bjorck, N., Gomes, C.P., Selman, B., Weinberger, K.Q. (2018). Understanding batch normalization. *Advances in Neural Information Processing Systems*, 31: 7694-7705.

- <https://doi.org/10.48550/arXiv.1806.02375>
- [49] Rawat, W., Wang, Z. (2017). Deep convolutional neural networks for image classification: A comprehensive review. *Neural Computation*, 29(9): 2352-2449. https://doi.org/10.1162/NECO_a_00990
- [50] Simonyan, K., Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*. <https://arxiv.org/abs/1409.1556>
- [51] Sangeetha, V., Prasad, K.J. (2006). Syntheses of novel derivatives of 2-acetylfuro [2, 3-] carbazoles, benzo [1, 2-b]-1, 4-thiazepino [2, 3-] carbazoles and 1-acetyloxycarbazole-2-carbaldehydes. *Indian J. Chem. - Sect. B Org. Med. Chem.*, 45(8): 1951-1954. <https://doi.org/10.1002/chin.200650130>
- [52] Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*. <http://arxiv.org/abs/1704.04861>
- [53] Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1251-1258. <https://doi.org/10.1109/CVPR.2017.195>
- [54] Yaman, O., Tuncer, T., Tasar, B. (2021). DES-Pat: A novel DES pattern-based propeller recognition method using underwater acoustical sounds. *Applied Acoustics*, 175: 107859. <https://doi.org/10.1016/j.apacoust.2020.107859>
- [55] Adegun, A., Viriri, S. (2021). Deep learning techniques for skin lesion analysis and melanoma cancer detection: A survey of state-of-the-art. *Artificial Intelligence Review*, 54: 811-841. <https://doi.org/10.1007/s10462-020-09865-y>