



Skin Lesion Classification Using CNN-based Transfer Learning Model

Kamil DIMILILER^{1,2,*} , Boran SEKEROGLU^{2,3} 

¹ Department of Electrical and Electronic Engineering, Near East University, Nicosia, North Cyprus, Mersin 10 Turkey

² Applied Artificial Intelligence Research Centre, Near East University, Nicosia, North Cyprus, Mersin 10 Turkey

³ Department of Computer Engineering, Near East University, Nicosia, North Cyprus, Mersin 10 Turkey

Highlights

- This paper focuses on the classification of skin lesions.
- A CNN-based transfer learning is implemented for the classification of skin lesions.
- The efficacy of transfer learning on skin lesion classification is studied.

Article Info

Received: 26 Jan 2022

Accepted: 09 Apr 2022

Keywords

Computer-aided
diagnosis
CNN
Skin lesions
Transfer learning
Cancer diagnosis

Abstract

The computer-aided diagnosis (CAD) and the analysis of skin lesions using deep learning models have become common in the last decade. The proposed CAD systems have considered various datasets and deep learning models. The transfer of knowledge from particular pre-trained models to others has also gained importance due to the efficient convergence and superior results. This study presents the design and implementation of a transfer learning model using Convolutional Neural Networks (CNN) with variable training epoch numbers to classify skin lesion images obtained by smartphones. The model is divided into the inner and external CNN models to train and transfer the knowledge, and the preprocessing and data augmentation are not applied. Several experiments are performed to classify cancerous and non-cancerous skin lesions and all skin lesion types provided in the dataset separately. The designed model increased the classification rates by 20% compared to the conventional CNN. The transfer learning model achieved 0.81, 0.88, and 0.86 mean recall, mean specificity, and mean accuracy in detecting cancerous lesions, and 0.83, 0.90, and 0.86 macro recall, macro precision, and macro F1 score in classifying six skin lesions. The obtained results show the efficacy of transfer learning in skin lesion diagnosis.

1. INTRODUCTION

Skin cancer is one of the most common malignancies worldwide. Melanoma and non-melanoma skin cancer incidence rates are increasing [1]. The reports show that more than 55,000 people die annually because of melanoma [2]. Therefore, early detection of skin cancer has vital importance in decreasing mortality rates and supporting therapy for the patients [3]. Due to different appearances of the lesions, there is a limited accuracy in diagnosing melanomas by an expert using visual and clinical inspection [4]. Asymmetry, border, color, and the diameter of the lesions are often used to assess the diagnosis [5]. Dermoscopy is used to increase the visual appearances of the lesions to provide detection of melanomas more accurately [6]. However, the experience of the experts is crucial for the consistent diagnosis of skin lesions.

In AI and deep learning, transfer learning is the use of the knowledge gained by a trained network in a different field in the same way that people transfer their experience from one area to an inexperienced one in real life. In deep learning, different models with different network architectures or training data were trained using millions of images, and the final weights, which are the knowledge of the neural network, were stored. These weights can be used to train other models for a particular problem. This process is known as transfer learning since the new models acquire the previous experiences of the pre-trained models.

* Corresponding author, e-mail: kamil.dimililer@neu.edu.tr

Meanwhile, several pre-trained networks such as Visual Geometry Group (VGG) networks (VGG16 and VGG19) [7], Residual Networks (ResNet) [8], InceptionV3 [9], and Densely Connected Convolutional Networks (i.e., DenseNet121) [10] were proposed and implemented based on CNN and effectively used in skin lesions classification tasks. Furthermore, these pre-trained neural networks have been publicly used by researchers that provide the transfer of gained and stored knowledge.

Therefore, the researchers have proposed several computer-aided diagnosis systems (CAD) to detect and classify skin lesions. However, the technological advancements in computers led to the use of deep learning for skin lesion classification, and accurate results were obtained. It is also possible to create, transfer and share the knowledge of a model to other models for a particular problem without using pre-trained networks. Since the knowledge of each neural network is stored, it could be used to train other models by partitioning the data. Tsiakmaki et al. [11] created models and implemented transfer learning using deep neural networks in order to predict student performance.

The primary aims of this study are to design a CNN-based transfer learning model for skin lesion diagnosis on a recent dataset that images acquired by smartphones and analyze the effect of training the data by transferring knowledge within the folds for a particular problem.

The rest of the paper introduces recent studies and the materials and methods of the study in Section 2 and Section 3. The designed transfer learning system is presented in detail in Section 4. Finally, the results and discussions part and the conclusion of the study are presented in Section 5 and Section 6, respectively.

2. LITERATURE REVIEW

Mahbod et al. [2] discussed the effect of transfer learning with multi-scale and multi-network systems to detect and classify skin lesions. In addition, the authors investigated the effect of the size of dermoscopic images based on pre-trained CNN with transfer learning. Khan et al. [12] used multiclass skin lesion detection and classification via Teledermatology. They proposed a hybrid system that fuses the binary images achieved or generated from a 16-layer CNN and segmentation based on the improved high dimension contrast transform. DenseNet201 is used in the classification phase using transfer learning.

Rodrigues et al. [13] proposed a new approach to classify skin lesions based on transfer learning, deep learning, and IoT. The authors proposed using transfer learning and deep learning in an IoT system to support the doctors in diagnosing common skin lesions using Convolutional Neural Networks as resource extractors. Several pre-trained networks and machine learning models were considered in their study. Hosny et al. [14] proposed a skin lesion classification system using Transfer Learning and AlexNet. The authors proposed that the parameters of the original model are used as initial values and initialize the weights of the last three replaced layers randomly.

Another study was performed by Zunair and Hamza [15] to detect melanoma using adversarial training and deep transfer learning. The authors proposed their research as two stages. In the first stage, the authors leveraged the inter-class mapping and synthesizing under-represented class samples from the over-represented ones using unpaired image-to-image translation. In contrast, a deep convolutional neural network has been trained in the second stage to classify skin lesions, considering the original training set with the newly synthesized under-represented class samples.

Afza et al. [16] proposed a three-step superpixel and deep learning for skin lesion classification. First, the authors applied contrast enhancement of the dermoscopy image set by fusing the local and global enhanced images and followed by image segmentation. ResNet-50 was applied in the research for the mapped images, and transfer learning was used for the learned features. The extracted features are optimized by the grasshopper optimization algorithm, followed by the Naïve Bayes classifier.

Singhal et al. [17] proposed a model based on the skin lesion classification using transfer learning. The model used four pre-trained networks, Inception v3, ResNet50, DenseNet201, and Inception ResNet v2. The authors trained their networks using seven different classes of skin lesions, and a comparative study was performed. Khatib et al. [18] proposed deep learning-based methods to diagnose skin lesions

automatically. In the first stage, a neural network is proposed to differentiate melanoma from benign nevus following the three pre-trained convolutional neural networks in the enumeration order. Then, the CNN architectures were fine-tuned to classify skin lesions using transfer learning.

Rahman and Ami [19] proposed a transfer learning-based approach for skin lesion classification using imbalanced data. The authors used ResNet, Xception, and DenseNet models to classify the skin lesions. Kondaveeti and Edupuganti [20] classified skin cancer images using transfer learning. The authors suggested a model that identifies the most common types of skin lesions. The authors used transfer learning utilizing multiple pre-trained models with class-weighted loss and augmentation techniques in the classification phase using ResNet50.

Jibhakate et al. [21] proposed skin lesion classification using deep learning and image processing. The authors compared the accuracies of convolutional neural networks with a transfer learning-based approach using Wide ResNet101, ResNet50, DenseNet121, and VGG19. Cauvery et al. [22] proposed a multiclass skin lesion classification system using a transfer learning-based convolutional neural network. They used to classify eight different skin lesions.

Islam et al. [23] proposed a deep learning model using image preprocessing. The authors compared the effect of normalization, data reduction, and data augmentation as preprocessing techniques with traditional deep learning. Bian et al. [24] applied multi-view filtered transfer learning on skin lesion classification. The authors proposed a multi-view filtered transfer learning network to represent discriminative information from various image views with a reasonable weighing strategy. In addition, the authors evaluated the importance of images that learn valuable knowledge by neglecting negative samples from the source domain and classified Melanoma and Seborrheic Keratosis.

Kumari and Sharma [25] reviewed recent research on skin lesion classification in their study. The authors discussed skin lesion classification tasks with Convolutional Neural Network according to their accuracy, precision, and other parameters using 13 types of CNN-based models.

The abovementioned skin lesions classification studies achieved accurate results generally for dermatologic images. However, the studies on a dataset of images acquired by smartphones are limited and could support places lacking medical devices and experts.

3. MATERIALS AND METHODS

3.1. Dataset

Contrary to the other skin lesion datasets consisting of dermoscopy images [26], the Dermatological Assistant Program (PAD) at the Federal University of Espirito Santo (UFES)-20 dataset [27] was released to support the CAD research to provide skin disease and cancer detection using the images captured by smartphones. This would yield effective CAD systems in rural areas with difficulties reaching medical equipment and experts [28].

The PAD-UFES-20 dataset consisted of 2298 images of 1373 patients with 1641 skin lesions. The skin lesions were classified as three skin diseases (Actinic Keratosis (ACK), Nevus (NEV, and Seborrheic Keratosis (SEK)) and three skin cancers (Basal Cell Carcinoma (BCC), Melanoma (MEL), and Squamous Cell Carcinoma (SCC)). Table 1 shows the number of images for each skin lesion type, and Figure 1 presents the sample images for each lesion type.

3.2. Convolutional Neural Networks

While the Convolutional Neural Network aims to simulate human beings' visual perception artificially and minimize preprocessing on images, it includes combined feature extraction and learning phases. A conventional CNN has three fundamental layers: convolution, pooling, and fully connected (dense) layers,

where the feature extraction process is performed in convolution and pooling layer, and the classification of the extracted features occurs in fully connected layers.

Table 1. Skin Lesion Types and Number of Images in PAD-UFES-20 Dataset

Diagnostic	Number of Images	Type
Basal Cell Carcinoma of skin (BCC)	845	Cancer
Actinic Keratosis (ACK)	730	Disease
Melanocytic Nevus of Skin (NEV)	244	Disease
Seborrheic Keratosis (SEK)	235	Disease
Squamous Cell Carcinoma (SCC)	192	Cancer
Malignant Melanoma (MEL)	52	Cancer

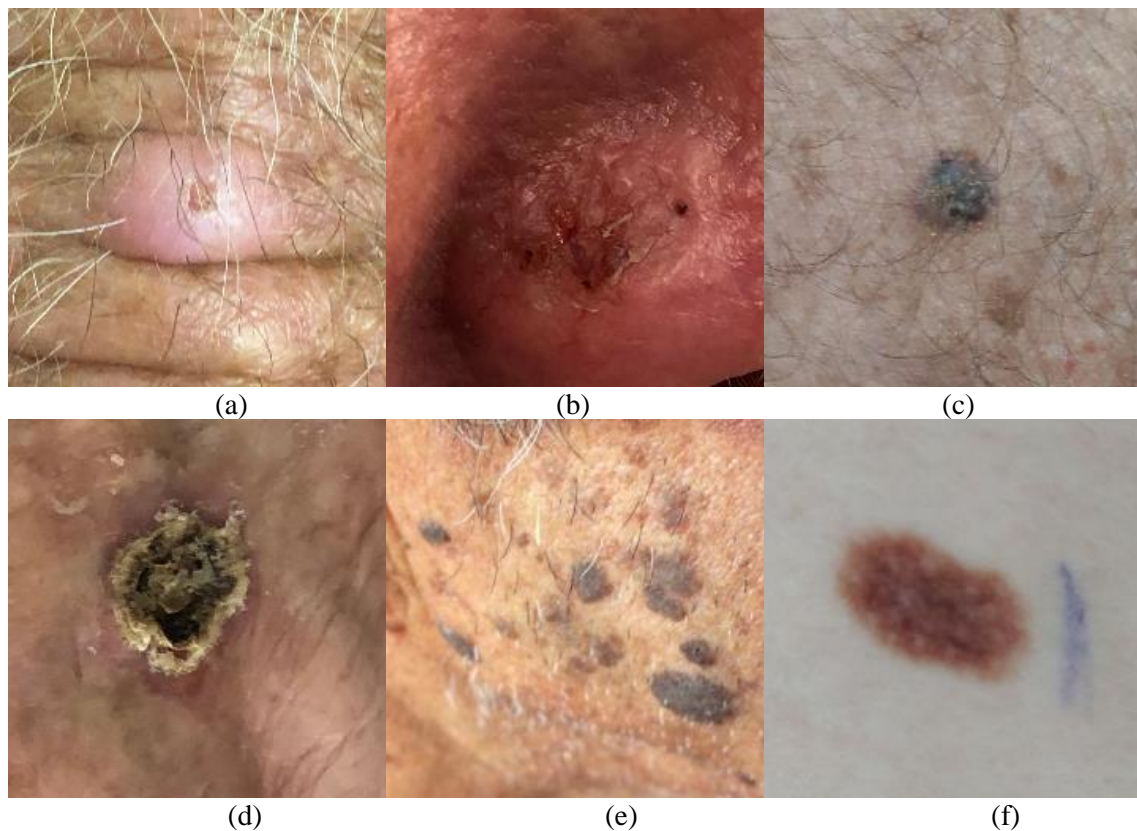


Figure 1. Sample skin lesion images, (a) SCC, (b) BCC, (c) MEL (d) ACK, (e) SEK, and (f) NEV

The feature extraction process is based on the pre-defined size filters applied within the convolutional layers. First, each mask detects the features on the corresponding images. The number of filters used on a single image directly affects the number of extracted features. Also, the size of the filters (i.e., 3x3, 5x5, etc.) defines the part of the image to be considered for the feature extraction instead of the whole image. Then, the activation function, which the Rectified Linear Unit (ReLU) is the common one, is applied to the obtained features to activate the most informative features non-linearly and to create a feature map for each image. The ReLU eliminates the negative values and activates only the positive values. Therefore, it provides faster convergence and decreases the models' computational cost. Applying activation functions to the features does not only provide effective feature extraction. In addition, it combines the learning and feature extraction process of particular features on spatial coordinates.

A pooling operation is applied to minimize the trainable parameters of the network by selecting the relevant points and reducing the dimension of created feature map based on the pooling type. This reduction operation has a vital effect on CNN by decreasing the size of the extracted features and also decreasing the computational cost of CNN. Pooling operation can be performed on pre-defined sizes using different pooling types (i.e., max-pooling, min-pooling, average-pooling). Min-pooling and average-pooling are based on the minimum and average values of features obtained by the filters, respectively. However, while the max-pooling considers the maximum value of the features that would store the most informative image component, it is the most common pooling type. Finally, the vectorized feature maps are fed to the fully connected layer.

The resultant classification process is performed within fully connected layers using the extracted features in convolutional layers.

3.3. Evaluation Metrics and Experiments

It is common for all classification studies to determine true-positive (TP), false-positive (FP), true-negative (TN), and false-negative (FN) values to measure the success of the model. True-positive and true-negative represent the correctly classified samples for the detected or not detected disease, respectively. Therefore, false-positive shows the wrong classification of the negative sample as positive by the model. On the contrary, false-negative shows the wrong classification of the positive sample as negative. Some of the standard evaluation metrics for binary (two-class) classification of balanced data are recall (sensitivity), specificity, and accuracy.

The recall is used to measure the ability of the models to predict the input samples with a disease correctly. The formula of the recall is given in Equation (1)

$$Recall = \frac{TP}{TP + FN}. \quad (1)$$

Contrary to the recall, specificity measures model capability to correctly predict the inputs samples without a disease. The formula of the specificity is given in Equation (2)

$$Precision = \frac{TN}{TN + FP}. \quad (2)$$

The accuracy is used to observe the general success of the models in using the balanced data. The formula of the accuracy is given in Equation (3).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \quad (3)$$

We considered the Precision, Recall, and F1 scores, which are common evaluation metrics used for balanced and imbalanced datasets in multinomial classification studies.

Precision measures true positives to all positives recognized by the model to measure the models' consistency while detecting the true positives. Equation (4) shows the formula of the precision

$$Precision = \frac{TP}{TP + FP}. \quad (4)$$

F1-score is the measure of models' general ability and is defined as the harmonic mean of precision and recall, as shown in Equation (5)

$$F1Score = 2x \frac{1}{\frac{1}{Precision} + \frac{1}{Recall}}. \quad (5)$$

In this study, the experiments were performed for two different tasks as binary and multinomial classification.

In the first experiments, a binary classification task was considered to classify the skin lesions as cancerous or non-cancerous. Squamous Cell Carcinoma, Malignant Melanoma, and Basal Cell Carcinoma of skin (BCC) images were bucketed into a cancerous class since the Actinic Keratosis, Melanocytic Nevus of Skin, and Seborrheic Keratosis are combined into non-cancerous class. Therefore, 1089 cancerous and 1209 non-cancerous images were considered in the experiments.

In multinomial experiments, the skin lesion classes were considered as they were provided in the dataset, and the experiments were performed to classify six types of skin lesions independently.

Conventional CNN and CNN-based transfer learning models were employed for both binary and multinomial experiments to compare the results obtained by direct learning and transfer learning. The evaluation of binary class experiments was performed using Sensitivity, Specificity, and Accuracy. In addition, macro-Precision, macro-Recall, and macro-F1 Scores were used to evaluate multinomial experiments.

3.4. Implementation of Transfer Learning-based CNN Model

This study used Convolutional Neural Networks with a lighter architecture to employ the transfer learning model for skin lesion detection. All models were trained using a five-fold cross-validation technique. In binary and multinomial classification, the complexity of learning in the network requires different architectures to achieve higher results.

Therefore, several CNN architectures were employed separately to determine the superior one for binary and multinomial experiments. Although CNN models achieved the highest results with the same number of convolutional layers for binary and multinomial experiments, differences occurred in the number of filters in convolutional layers and the number of neurons in the dense layer.

A CNN model that included two convolutional layers with 32 and 16 filters and three fully connected layers with 64, 16, and 2 neurons, respectively, was determined as superior architecture for binary classification experiments (CNN Model A).

For multinomial experiments, the architecture of the final CNN model included two convolutional layers with 64 and 32 filters and three dense layers with 128, 32, and 6 neurons, respectively (CNN Model B). Table 2 summarizes the experiments, types, and considered models.

Table 2. Properties of experiments

Experiment Name	Type	Model
Binary-CNN	Binary Classification	CNN Model A
Binary-TL	Binary Classification	CNN Model A
Multi-CNN	Multinomial Classification	CNN Model B
Multi-TL	Multinomial Classification	CNN Model B

Commonly for all experiments, all filters were implemented with a size of 3x3, and 2x2 maximum pooling was applied to each layer. In addition, 10% of drop out was considered for each layer to prevent overfitting during the convergence. The 'Adam' optimizer was considered during the weight updating.

The conventional CNN models were trained for 100 epochs for all experiments, however, the transfer learning models were trained using variable epoch numbers. The epoch numbers were started from 100 and

decreased by 20 for each fold. Therefore, the last fold of each experiment for the transfer learning model was 20.

The mini-batch was used as 32, and the color input images were fed to the models in 100x100 dimensions. Figure 2 shows the architecture of the CNN model considered in the multinomial experiments. Table 3 presents the common parameters of the CNN models considered in the experiments.

Table 3. Common parameters of CNN models

Parameter	Value
Filter Size	3x3
Drop-out for each layer	10%
Optimizer	Adam
Mini-batch size	32
Input image size	100x100
# of Training epochs (without TL)	100
# of Training epochs (with TL)	variable

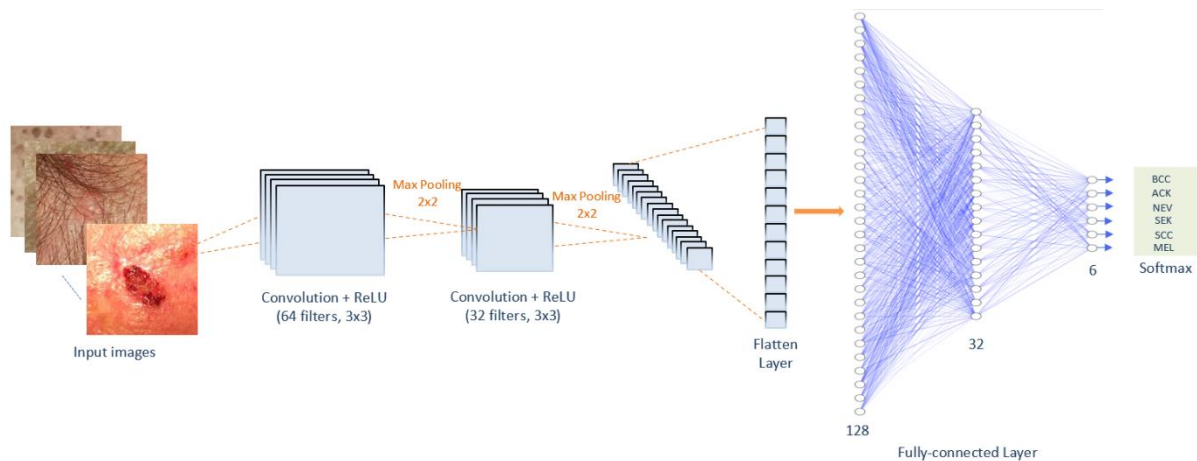


Figure 2. Considered CNN architecture of models

The transfer learning model was obtained by transferring the weights of CNN models obtained from the training folds in the five-fold cross-validation to the next training, which was applied by including the fold used as the test fold in the previous training.

First, the dataset was split into 5x5 folds, as external and inner, using five-fold cross-validation, and five external models with five inner transfer models were created. While the inner models are single-run five-fold cross-validation, each external model was linked together to include five internal models for each.

In this way, the knowledge of previous inner models was transferred to the next model with added new images, and the training of the next model was performed. This yielded the training of each external model using the pre-trained images with the different combinations of training data for each inner model.

In order to change the training and testing data dynamically, the internal test folds were altered according to the standard procedure at each training. In the training of internal models, the test fold of the corresponding external model was not included in the training data, and only the training folds were considered. In this way, the external folds were used as the untrained test image source of the models. Then, the internal models transferred the weights of each finalized training to the next training. Thus, while the inner folds were rapidly transferring weight between themselves, no weight transfer was made between the external models, only the data was changed to provide all images to be considered both during the train and test phase.

The evaluation of the model was performed by calculating the mean of the abovementioned metrics for the external test folds results.

The models were trained according to the variable epoch number since the excess epoch number may cause overfitting in each increment of the inner models because of the knowledge provided by transfer learning. While Inner Model 1 was trained with 100 epochs, the number of training epochs was shortened by 20 epochs in each inner model increment. Therefore, Inner Model 1, Inner Model 2, Inner Model 3, Inner Model 4, and Inner Model 5 were trained with 100, 80, 60, 40, and 20 epochs, respectively. Figure 3 shows the block diagram of the transfer learning model in detail.

5. EXPERIMENTAL RESULTS AND DISCUSSIONS

This section presents the results obtained for binary and multinomial classification tasks and the discussions on the results. As mentioned above, the experiments were performed for binary and multinomial classification and conventional CNN without transfer learning, and the designed CNN-based transfer learning was considered in the experiments. Furthermore, all experiments were performed using the five-fold cross-validation technique.

5.1. Experimental Results

The CNN model with lighter architecture and without transfer learning in binary experiments (cancerous or non-cancerous) obtained 0.70, 0.66, and 0.68 mean recall, mean specificity, and mean accuracy. In the same experiment, the transfer learning model achieved 0.81, 0.88, and 0.86 mean recall, mean specificity, and mean accuracy and provided 0.114, 0.226, and 0.184 improvements.

As mentioned above, the CNN models with different parameters were considered in multinomial experiments. As a result, the CNN model without transfer learning obtained 0.676, 0.643, and 0.701 macro F1 Score, macro Recall, and macro Precision. On the other hand, contrary to the lower results obtained without transfer learning, the transfer learning model achieved 0.865, 0.839, and 0.9030 macro F1 Score, macro Recall, and macro Precision. Table 4 shows the obtained results for binary and multinomial classification experiments.

5.2. Grad-Cam Analysis of the Results

Gradient-weighted Class Activation Mapping (Grad-CAM) [29] visualizes the points at which the network is activated for target classes. It uses gradients that flow to the last convolution layer of the convolutional neural network and understands the importance of each neuron for a target decision.

Even though we could not determine how the network learns these points, it provides valuable information about the network outputs to observe the points that affect the decisions.

In this study, we performed Grad-Cam analysis to analyze the effect of transfer learning on the untrained data. The Grad-Cam results of the first inner model of the multinomial experiments were compared to the last inner model. The first inner model was the initial model and did not include any transferred knowledge, however, the last inner model gained knowledge by transferring the previous weights.

The analysis of Grad-Cam results showed that the gained knowledge and the improved classification rates caused transfer learning models to focus the features more particularly than the initial model. The improved recognition rates provided by transfer learning eliminated the activation of irrelevant features on the spatial coordinates, allowing the researcher to better understand the diagnostic decisions of artificial intelligence models.

Figure 4 demonstrates the difference between the initial inner model activations and the final inner model activations using the small values of the gradients. The presented images show how the irrelevant regions could activate the initial model. In contrast, the final inner model minimized the irrelevant regions and focused particularly on the lesion regions within the images.

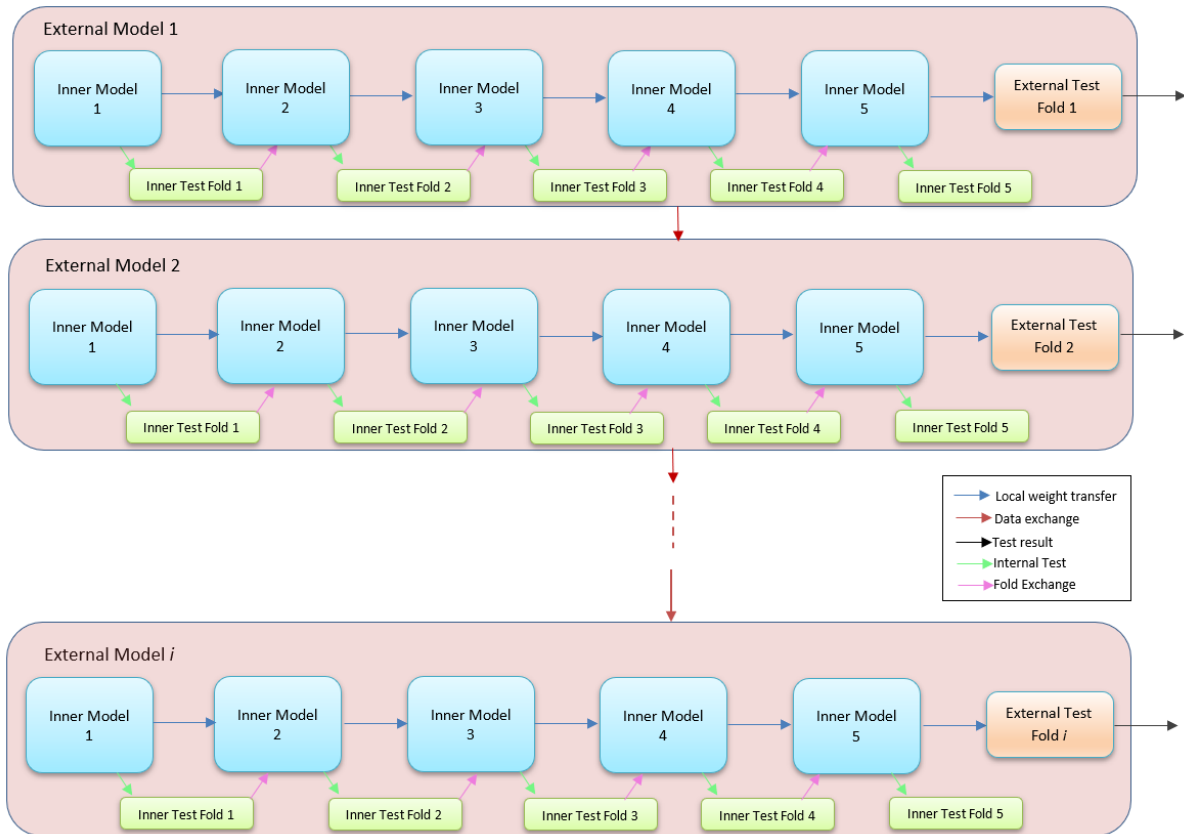


Figure 3. General block diagram of transfer learning model

Table 4. Results of the study

Binary Classification Experiments			
Model	Mean Recall	Mean Specificity	Mean Accuracy
Binary-CNN	0.70	0.66	0.68
Binary-TL	0.81	0.88	0.86
Multinomial Classification Experiments			
Model	Macro Recall	Macro Precision	Macro F1 Score
Multi-CNN	0.64	0.70	0.67
Multi-TL	0.83	0.90	0.86

5.3. Discussions

It is common knowledge that obtaining pre-trained knowledge during the new training is beneficial for the neural networks for particular cases [30]. Therefore, even though various transfer learning studies have been performed, the researchers generally focus on transferring knowledge from the pre-trained deep models such as Res-Net models instead of the lighter CNN model.

Both binary and multinomial classification experiments showed that obtaining and transferring knowledge through the training folds increased the rates of skin lesion diagnosis up to 21% compared to the direct learning model. In addition, the Grad-Cam analysis of the results demonstrated how the ability of the models was improved and focused on the region of interest while transferring knowledge and increasing the classification accuracy.

The results of this study were obtained without any preprocessing and data augmentation procedure, therefore, the effect of transfer learning for skin lesion diagnosis was demonstrated using the raw images.

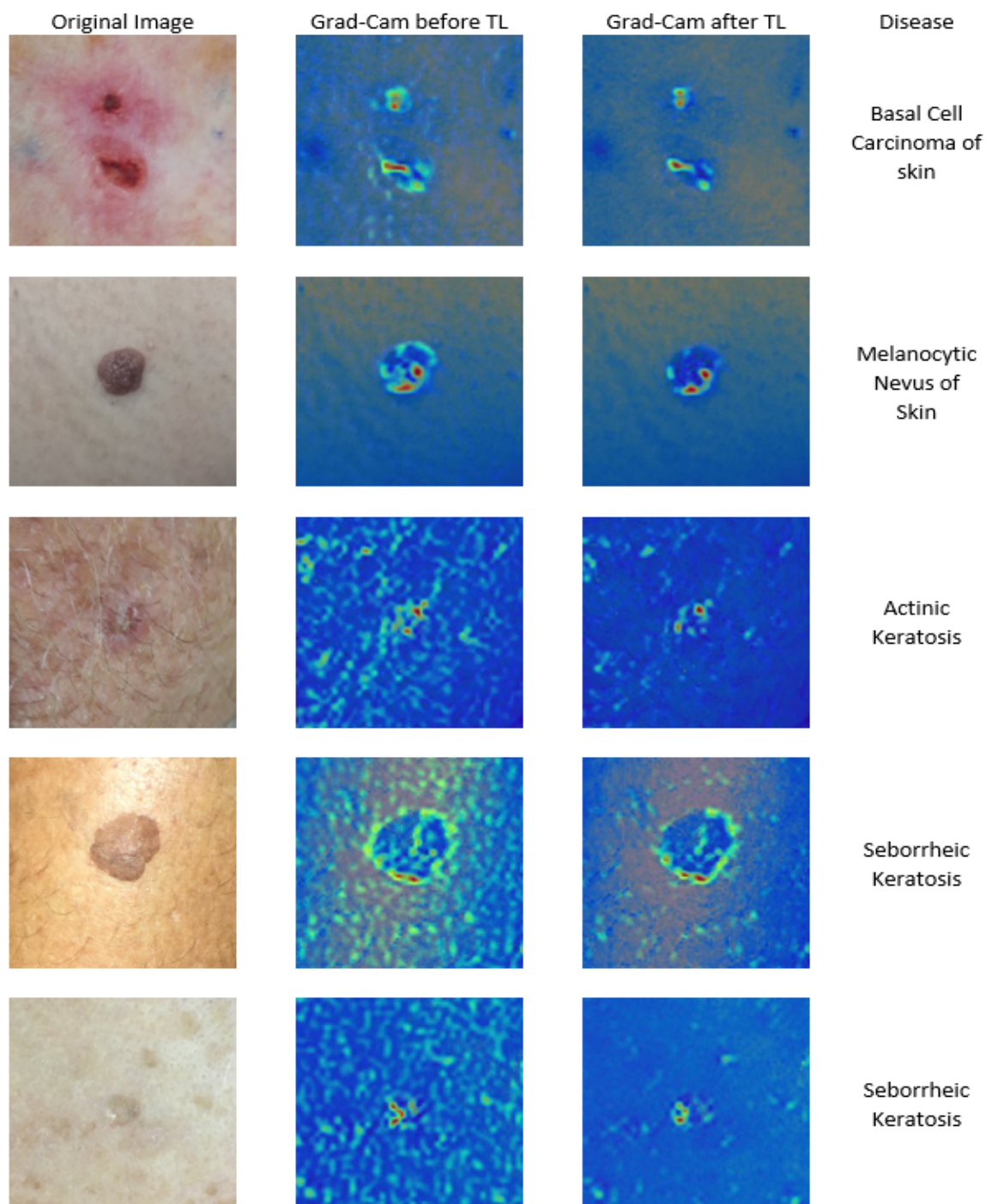


Figure 4. Grad-Cam analysis of the results (small values)

Table 5. Performance comparison of the suggested model with the models in literature

Relevant Research	Classification Accuracy
Krohling et al. [31]	85%
Pacheco and Krohling [32]	80%
Karthik et al. [33]	84.7%
Khan et al. [34]	76%
Khan et al. [34] (with clinical features)	78%
Our study	86%

Table 5 shows the comparison of relevant research papers and their accuracies. Krohling et al. [31] performed experiments similar to the binary classification experiments of this study with the addition of clinical information of the patients. Their study achieved 85% of balanced accuracy.

Pacheco and Krohling [32] investigated the effect of using patients' clinical information in the diagnosis of skin lesions. Their study showed that considering clinical information with the skin lesion images significantly increased the classification rates. The precision, recall, and F1 scores were achieved as 0.80, 0.78, and 0.79, respectively, for the multinomial experiments considered in this study.

Karthik et al. [33] proposed an efficient channel attention-based convolutional neural network for skin disease classification. The proposed network is used with 16 M parameters to classify the disease. The proposed research was performed on four classes: acne, actinic keratosis (AK), melanoma, and psoriasis with an overall testing accuracy of 84.70%.

Khan et. al. [34] performed the experiments considering skin lesions and skin lesions with clinical features. The authors proposed a Remote Diagnosis and Triaging Model to detect skin cancer using EfficientNet and Extreme Gradient Boosting algorithms. The authors achieved an accuracy of 76% and 78% using clinical features.

The recent research and the obtained results showed that balancing the data and using clinical information of the patients are valuable to increase the classification rates. However, transfer learning improved the results significantly, even though the preprocessing, data augmentation and clinical information were not considered in the experiments.

Proper preprocessing, data augmentation, and feeding patient information to the CNN model by combining the transfer learning model require further investigation to improve the recognition rates.

6. CONCLUSION

The classification of skin lesions has great importance in providing early diagnosis, treatment, and precautions. Since the deep learning models have strong abilities to classify images for a particular problem, they were implemented, proposed, and performed this task accurately, particularly within the last decade. In addition, the use of transfer learning improves models' convergence and therefore provides a more accurate classification of the labels.

In this paper, we designed a transfer learning model using the conventional CNNs and performed binary (cancer vs. non-cancerous) and multinomial (six lesions) classification on the recently released dataset without applying preprocessing, data augmentation, and without considering the patients' information during the training.

The obtained results demonstrated the significant effect of transfer learning by increasing the classification rates by more than 20% compared to conventional models. The CNN-based transfer learning model achieved 0.86, and 0.86 mean accuracy and F1 score for binary and multinomial classification of skin lesions.

The efficacy of transfer learning is demonstrated once more, and its' effect on the skin lesion classification is shown.

Our future work will focus on improving the obtained results by preprocessing images and increasing the number of images using data augmentation.

CONFLICTS OF INTEREST

No conflict of interest was declared by the authors.

REFERENCES

- [1] Leiter, U., Eigentler, T., Garbe, C., “Epidemiology of skin cancer”, *Advances in Experimental Medicine and Biology*, 810: 120-40, (2014).
- [2] Mahbod, A., Schaefer, G., Wang, C., Dorffner, G., Ecker, R., Ellinger, I., “Transfer learning using a multi-scale and multi-network ensemble for skin lesion classification”, *Computer Methods and Programs in Biomedicine*, 193(2020): 105475, (2020).
- [3] Schadendorf, D., van Akkooi, A., Berking, C., Griewank, K., Gutzmer, R., Hauschild, A., Stang, A., Roesch, A., Ugurel, S., “Melanoma”, *Lancet*, 392(10151): 971–984, (2018).
- [4] Brinker, T., Hekler, A., Utikal, J., Grabe, N., Schadendorf, D., Klode, J., Berking, C., Steeb, T., Enk, A., Von Kalle, C. “Skin cancer classification using convolutional neural networks: systematic review”, *Journal of Medical Internet Research*, 20(10), (2018).
- [5] Stolz, W., Riemann, A., Cognetta, A., Pillet, L., Abmayr, W., Holzel, D., Bilek, P., Nachbar, F., Landthaler, M., Braun-Falco, O, “ABCD Rule of dermatoscopy: a new practical method for early recognition of malignant melanoma”, *European Journal of Dermatology*, 4: 521–527, (1994).
- [6] Tromme, I., Sacré, L., Hammouch, F., Legrand, C., Marot, L., Vereecken, P., Theate, I., Van Eeckhout, P., Richez, P., Baurain, J., Thomas, L., Speybroeck, N., on behalf of the DEPIMELA study group, “Availability of digital dermoscopy in daily practice dramatically reduces the number of excised melanocytic lesions: results from an observational study”, *British Journal of Dermatology*, 167(4): 778–786, (2012).
- [7] Simonyan, K., Zisserman, A., “Very deep convolutional networks for large-scale image recognition”, *Proceeding of the 2015 International Conference on Learning Representations*, (2015).
- [8] He, K., Zhang, X., Ren, S., Sun, J., “Deep residual learning for image recognition”, *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, 770–778, (2016).
- [9] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., “Rethinking the inception architecture for computer vision”, *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, 2818–2826, (2016).
- [10] Huang, G., Liu, Z., van der Maaten, L., Weinberger K., “Densely connected convolutional networks”, *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*, 2261–2269, (2017).
- [11] Tsiakmaki, M., Kostopoulos, G., Kotsiantis, S., Ragos, O., “Transfer learning from deep neural networks for predicting student performance”, *Applied Sciences*, 10(6): 2145, (2020).
- [12] Khan, M.A., Muhammad, K., Sharif, M., Akram, T., de Albuquerque, V.H.C., “Multi-class skin lesion detection and classification via teledermatology”, *IEEE Journal of Biomedical and Health Informatics*, 25(12): 4267-4275, (2021).
- [13] Rodrigues, D. D. A., Ivo, R.F., Satapathy, S.C., Hemanth, J., Filho, P.P.R., “A new approach for classification skin lesion based on transfer learning, deep learning, and IoT system”, *Pattern Recognition Letters*, 136: 8–15, (2020).
- [14] Hosny, K., Kassem, M.A., Fouad, M.M., “Classification of skin lesions using transfer learning and augmentation with Alex-net”, *PLoS ONE*, 14(5), (2019).

- [15] Zunair, H., Hamza, A.B., “Melanoma detection using adversarial training and deep transfer learning”, *Physics in Medicine Biology*, 65(15): 135005, (2020).
- [16] Afza, F., Sharif, M., Mittal, M., Khan, M.A., Hemanth, D.J., “A hierarchical three-step superpixels and deep learning framework for skin lesion classification”, *Methods*, (2021).
- [17] Singhal, A., Shukla, R., Kankar, P., Dubey, S., Singh, S., Pachori, R., “Comparing the capabilities of transfer learning models to detect skin lesion in humans”, *Proceedings of the 2020 Institution of Mechanical Engineers Part H Journal of Engineering in Medicine*, 234(10): 1083–1093, (2020).
- [18] El-Khatib, H., Popescu, D., Ichim, L., “Deep learning-based methods for automatic diagnosis of skin lesions”, *Sensors*, 20(6): 1753, (2020).
- [19] Rahman, Z., Ami, A., “A transfer learning based approach for skin lesion classification from imbalanced data”, *Proceedings of the 2020 11th International Conference on Electrical and Computer Engineering*, 65–68, (2020).
- [20] Kondaveeti, H., Edupuganti, P., “Skin cancer classification using transfer learning”, *Proceedings of the 2020 IEEE International Conference on Advent Trends in Multidisciplinary Research and Innovation*, 1–4, (2020).
- [21] Jibhakate, A., Parnerkar, P., Mondal, S., Bharambe, V., Mantri, S., “Skin lesion classification using deep learning and image processing”, *Proceedings of the 2020 3rd International Conference on Intelligent Sustainable Systems*, 333–340, (2020).
- [22] Cauvery, K., Siddalingaswamy, P., Pathan, S., D'souza, N., “A multiclass skin lesion classification approach using transfer learning based convolutional neural network”, *Proceedings of the 2021 Seventh International Conference on Bio Signals, Images, and Instrumentation*, 1–6, (2021).
- [23] Islam, M.K., Ali, M.S., Ali, M.M., Haque, M.F., Das, A., Hossain, M., Duranta, D., Rahman, M., “Melanoma Skin Lesions Classification using Deep Convolutional Neural Network with Transfer Learning”, *Proceedings of the 2021 1st International Conference on Artificial Intelligence and Data Analytics*, 48–53, (2021).
- [24] Bian, J., Zhang, S., Wang, S., Zhang, J., Guo, J., “Skin lesion classification by multi-view filtered transfer learning”, *IEEE Access*, 9: 66052–66061, (2021).
- [25] Kumari, A., Sharma, N., “A review on convolutional neural networks for skin lesion classification”, *Proceedings of the 2021 2nd International Conference on Secure Cyber Computing and Communications*, 186–191, (2021).
- [26] Tschandl, P., Rosendahl, C., Kittler, H., “The HAM10000 dataset: A large collection of multi-source dermatoscopic images of common pigmented skin lesions”, *Scientific Data*, 2018, 5(1): 1–9, (2018).
- [27] Pacheco, A., Lima, G., Salomão, A., B.Krohling., Biral, I., de Angelo, G., Alves Jr, F., Esgario, J., Simora, A., Castro, P., Rodrigues, F., Frasson, P., Krohling, R., H.Knidel., Santos, M., do Espírito Santo, R., Macedo, T., Canuto, T., de Barros, L., “PAD-UFES-20: A skin lesion dataset composed of patient data and clinical images collected from smartphones”, *Data in Brief*, 32: 106221, (2020).
- [28] Farajnia, S., “Overcoming the data gap for the remote diagnosis of skin cancer”, *Patterns (NY)*, 1(7): 100117, (2020).

- [29] Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D., “Grad-CAM: visual explanations from deep networks via gradient-based localization”, Proceedings of the 2017 IEEE International Conference on Computer Vision, 618–626, (2017).
- [30] Weiss, K., Khoshgoftaar, T., Wang, D., “A survey of transfer learning”, Journal of Big Data, 3(1): 1-40, (2016).
- [31] Krohling, B., Castro, P.C., Pacheco, A.C., Krohling, R., “A smartphone based application for skin cancer classification using deep learning with clinical images and lesion information”, arXiv preprint arXiv:2104.14353, (2021).
- [32] Pacheco, A., Krohling, R., “The impact of patient clinical information on automated skin cancer detection”, Computers in Biology and Medicine, 116: 103545, (2020).
- [33] Karthik, R., Vaichole, T. S., Kulkarni, S. K., Yadav, O., Khan, F., “Eff2Net: An efficient channel attention-based convolutional neural network for skin disease classification”, Biomedical Signal Processing and Control, 73: 103406, (2022).
- [34] Khan, I. U., Aslam, N., Anwar, T., Aljameel, S. S., Ullah, M., Khan, R., Akhtar, N., “Remote Diagnosis and Triaging Model for Skin Cancer Using EfficientNet and Extreme Gradient Boosting”, Complexity, 2021: 5591614, (2021).