

# A CNN Framework for Classification of Melanoma and Benign Lesions on Dermatoscopic Skin Images

Erdal Özbay\*

Department of Computer Engineering, Firat University, Elazig, Turkey - 23119  
Email: erdalozbay@firat.edu.tr

Feyza Altunbey Özbay

Department of Software Engineering, Firat University, Elazig, Turkey - 23119  
Email: faltunbey@firat.edu.tr

---

## ABSTRACT

---

Melanoma is the most common type of skin cancer due to a genetic predisposition. In recent years, it has been determined that the number of different types of skin cancer has increased worldwide and caused a large number of deaths. Some skin cancers, such as melanoma and its derivatives, can be prevented, but early and accurate diagnosis is very important for treatment. Image processing techniques in medical applications are frequently used in the diagnosis, follow-up, and treatment processes of skin cancer. However, manual control of medical images is laborious and time-consuming and is vulnerable to expert errors in the interpretation of images. Developing a safe and autonomous classification system for medical applications is a fundamental need. In this study, a CNN-based deep learning framework has been developed in which the HAM10000 dataset, a dermatoscopic clinical skin image collection, has been classified for skin cancer detection. Classification preprocessing using contrast limited adaptive histogram equalization is demonstrated by the accuracy results that improve the recognition of subtle features of class labels. A 45-layer model is proposed for classification. With this developed model, an accuracy rate of 99.69% has been achieved. The results show that the proposed model achieves high accuracies and F-measures with low false-negative compared to known classifiers. This CNN model showed the best two-level performance classifying melanoma and benign cases as nevi and non-nevi. It has emphasized that skin cancer can be detected early with the proposed model and can contribute to the execution of the treatment process.

Keywords -Classification, CNN, Deep learning, Image processing, Melanoma.

---

Date of Submission: Sep 30, 2021

Date of Acceptance: Oct 18, 2021

---

## I. INTRODUCTION

The skin is the organ with the largest volume that balances the temperature of the human body and protects it against infections and light. In addition to keeping the body temperature in balance, it is the body's water and fat storage center. The skin is a barrier that protects people against external environments. The skin, which consists of three layers called the dermis, stratum basale, and stratum corneum carries out the elimination of various harmful substances in the body through sweat. The biggest danger that the skin can face in the body is skin cancer [1].

There are three basic types of skin cancers, which can occur in different colors and shapes; basal cell, squamous cell, and melanoma. Skin cancer, which is one of the most common types of cancer, can occur due to many reasons such as various infections, allergies, and excessive exposure to the sun. In this regard, skin cancer can also be seen in parts of the body that are not exposed to sunlight, apart from parts of the face, hands, and neck that are more exposed to Ultraviolet (UV) rays. The cells that nourish the skin age and die daily. Skin cells grow and divide to form new cells, along with the basic components that make up the skin so that new cells replace the dead. In some cases, this systematic order is disrupted, cells die

even though they do not age, and new cells are formed when they are not needed. The extra accumulation of skin cells in the body causes the formation of a tissue mass called a tumor. In this regard, abnormal and uncontrolled proliferation of skin cells in the body leads to skin cancers [2].

Melanoma, known as cutaneous malignant melanoma (MM), is the most noxious type of skin cancer. Melanoma has been the 5th most common cancer type with the number of cases exceeding one million every year since 2018. Approximately 5000 men and 2500 women who were diagnosed with melanoma died in 2019. Melanoma is a deadly skin cancer that starts in the melanocyte cells that give color to human skin. Most MM skin cancer cells usually continue to produce melanin. Therefore, the tumor is usually brown or black in these patients. Although MM is a common type of skin cancer due to genetic predisposition, it has special importance because it is the deadliest skin cancer [3]. People with one or more first-degree relatives with melanoma have a higher risk. About 10% of people with melanoma have a family history of the same disease. The reasons for the increased risk may consist of similar family life factors such as the presence of sensitive light-skinned family members and sun exposure. Gene mutations have been observed in an

average of 30% of family members with a high rate of melanoma. In this regard, gene mutation may be a factor in the risk of melanoma. MM accounts for about 5% of skin cancers, although MM causes an average of 85% of skin cancer-related deaths [4].

Melanoma, one of the types of skin cancer, has a low detection rate, although it causes most of the deaths that occur. As with other types of cancer, early detection of skin cancer increases the success rate of treatment. Conversely, while undiagnosed, cancer cells can spread to other parts of the body, resulting in irreversible consequences [5]. The main problem in the detection of melanoma is that they have similar features to melanocytic nevi (MN). In this respect, distinguishing benign moles and melanomas in the first growth phase is a challenging process for dermatologists. Dermatologists adhere to several rules for classifying and distinguishing between them. For this purpose, various computer-aided algorithms with low complexity methods have been developed and run on smartphones and tablets to help experts. Even these methods, which require the use of various hardware and have a complex algorithm structure, ultimately need to make a professional decision [6]. In this case, cancer identification can be made simpler with a system to be developed, and experts can be helped to interpret the diagnosis if they are undecided. With the help of these systems, expert mistakes that occur with manual methods will be reduced, the diagnosis time will be shortened, the workload of radiologists will decrease while the quality of melanoma diagnosis will increase [7-9].

Today, with the rapid development of technology, there is a great increase in storing data in databases, especially in the field of medicine. In recent years, many methods have been used to classify these data on various cancer diseases. Classifying and interpreting large amounts of stored data with known methods is a laborious and rather difficult process. The learning process is provided by using big data in databases with a model developed with deep learning architectures that have been popular in the last few years. The model trained with these data passes the testing process and is subjected to the calculation of performance criteria [10].

Our study aims to implement an autonomous melanoma skin cancer classifier that does not require manual intervention and a parameter setting. In this system, skin diseases are classified by training the data used with a developed model. In this study, a new 45-layer model was developed for deep learning. It is aimed to diagnose skin cancer through clinic dermatoscopic images with the proposed model. A dataset named MNIST: HAM10000 has been used which contains 10015 dermatoscopic images in total [11]. The proposed model is trained with 8012 image data. Then, 2003 data have been used for the test phase. So, 20% of the data set has been used for testing and 80% for training. The developed model accurately predicted about 1998 of the test data used. In this article, deep learning performance trained by applying state-of-the-art transfer learning with a developed model is evaluated in melanoma detection. The result of this experiment showed that the model applied in the

developed method has a high accuracy rate in the diagnosis of skin cancer.

The next sections of the article are presented in the following order; Section 2 summarizes the related works using convolutional neural networks for classification. Section 3 describes proposed methodologies for the classification of melanoma and benign lesions. Section 4 includes experimental results and discussions. Lastly, Section 5 describes the main conclusions of the work.

## II. RELATED WORKS

Professional detection and classification of benign and malignant melanomas is a challenge, even for experienced dermatologists. The main reason for this is that MN has similar characteristics to MM. Therefore, dermatologists want to follow a certain rule for this job. Some accepted rules in dermatology guide specialists in this task. One of them is the ABCDE (Asymmetry, Border irregularity, Color patterns, Diameter, and Evolving) rule [12]. The traditional skin analysis rule accepted by dermatologists is shown in Fig.1.



Figure 1. Traditional clinical analysis rule (ABCDE).

A MM is irregularly asymmetrical, whereas a MN is usually symmetrical. A MM is often irregular and difficult to define, whereas a MN has a smooth structure with even borders. A MM can have six different colors that are irregularly dispersed, whereas a MN usually has a single or two regularly recurring hues. One rule is for diameter. At the diagnostic stage, many MMs were found to be larger than 6 mm in size. Another rule is to evolve. Although MM progresses in color, size, and shape from diagnosis, MN is usually stable [13].

Automatic classification systems, which are obtained by developing prediction models on medical images, are generally exposed to low-capacity data usage. Besides, the presence of unwanted objects and homogeneity in the images that make up the training datasets cause some pre-processes such as segmentation of the region [14]. In addition, various methods of feature extraction can be used to improve the success rate in classification [15]. Although the extraction of some features has been performed from melanoma images to develop classification models in many studies, these approaches require a long process and expert need. In this respect, the image preprocessing used in some methods may cause the loss of some information on the data or the formation of mistakes that may affect the classification success rate. Especially a few years ago, poor segmentation of a skin lesion was performed with low accuracy in traditional classification methods [16, 17]. To cope with these classification constraints, in recent years, some deep learning models have been developed

that can distinguish significant features between image classes by automatically learning determinants. In this respect, deep convolutional neural networks (CNN) are a different type of artificial neural network that has been able to produce strong results in general, especially in different image processing and medical imaging applications in recent years [18].

In a recent study, melanoma detection was performed based on deep neural networks and a dermatologist-level classification was proposed. The dataset images used were directly trained end-to-end with a single CNN and melanoma was classified. The efficacy of this study was confirmed by biopsy of dermatoscopic images approved by a large number of dermatologists [19].

In many proposed studies, deep learning, which does not require an image preprocessing, has been used to successfully solve complex classification and segmentation problems. The model architecture of these networks usually includes convolutional layers that are filtered by extracting the defining features of the images so that different types of lesions can be learned. For example, CNNs that perform the learning of the crucial key features of dementia cases using different modality images have shown high performance [20].

CNNs are frequently used in popular topics such as MR-images fusion [21], breast cancer [22], tumor diagnosis [23], lesion classification [24], and segmentation in recent years, due to their high precision and ability to process large amounts of data in a short time using GPUs. In addition, neural networks fed with extra-privileged information have been used to perform strain reconstruction in ultrasound elastography with deep learning models. Moreover, studies have been performed to detect blood flow rate from angiography images and to detect vessel borders with intracoronary images [25, 26].

Satisfactory performance results have been obtained in data analysis using CNN, as datasets containing medical images becoming publicly accessible. For this, images are first divided into several smaller superpixels, and operators are applied to these pixels. Although important studies have been carried out on the classification of skin lesions, this subject is open to development. These studies usually consist of two phases; segmentation of crucial features with deep networks and then estimating them. In this regard, the process focuses on two separate classes, and different skin classes are considered in the same class and not classified [27].

When the studies carried out in this field are examined so far, it is concluded that the use of CNN models increases the efficiency of the diagnostic system. However, in some studies, these methods have been supported by various optimization algorithms [28]. In some studies, the whale optimization algorithm has been used to improve the

convolutional neural network. This algorithm is inspired by the bubble technique that humpback whales use to catch their prey [29].

### III. PROPOSED METHODOLOGIES

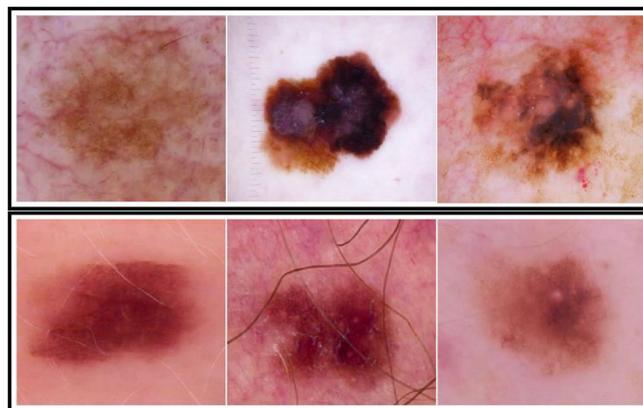
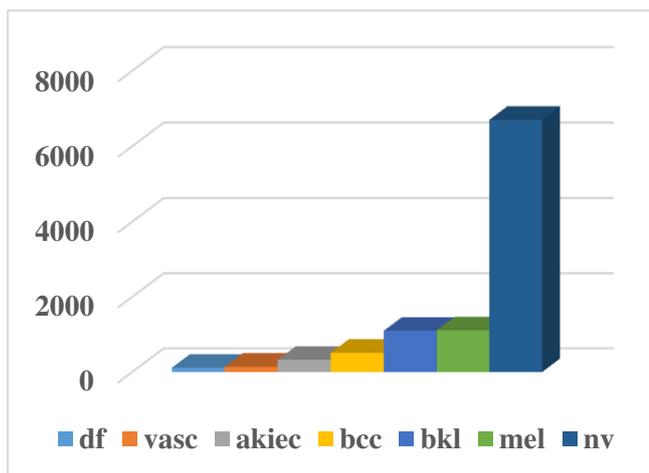


Figure 2. Examples of dataset images for melanoma and benign nevi.

This study aims to determine a total of 10015 dermatoscopic skin images. Of these, 8012 have been used for training the model and the remaining 2003 for testing. In this article, the model developed for the classification of melanoma and benign lesions and CNN models are analyzed. The proposed model showed that the data in the skin cancer type have been classified with high accuracy and the results obtained showed that the performance efficiency of the developed model has high. Some examples in the dataset for both classes are shown in Fig.2.

#### 3.1. DATASET

The HAM10000 dataset, whose analysis is publicly available, has been used in the training and testing stages of the developed CNN model and in obtaining experimental results [11]. This dataset is frequently used as a benchmark for comparing manual and machine classification and deep classification networks approaches. The dataset has 10015 dermatoscopic images obtained over 20 years from the dermatology department of the Medical University of Vienna and the skin cancer practice of Cliff Rosendahl. The dataset was collected in seven different classes, the majority of which were the nevi (nv) class. Apart from the benign lesion, melanoma and its potential derivatives include classes; dermatofibroma (df), vascular skin (vasc), actinic keratosis (akiec), basal cell carcinoma (bcc), benign keratosis (bkl), and melanoma (mel).



As shown in Fig.3, it is a significant disadvantage that this dataset has unequal numbers of different classes. This unbalanced distribution situation consists of data belonging to 6705 nevi classes, which are considered harmless, whereas the remaining data consists of 3310. When the dataset is evaluated in this respect, it causes specialization in this class due to the abundance of images with similar features to benign nevi, while attention should be paid to test analyzes in this class due to the lesser number of vascular skin images. As a result, in order to balance this irregular distribution of the data set, other data outside the benign nevi class at the training stage have been grouped under the melanoma class. The images in the data set are 600x450 pixels and they have been saved in .jpg format. 80% of the images in the whole data set have been used for training and 20% for testing.

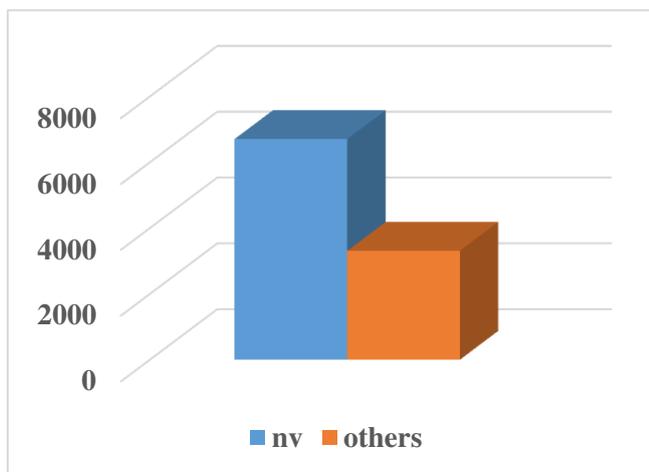


Figure 3. HAM10000 dataset, original and two-class distribution.

### 3.2. Digital Image Processing

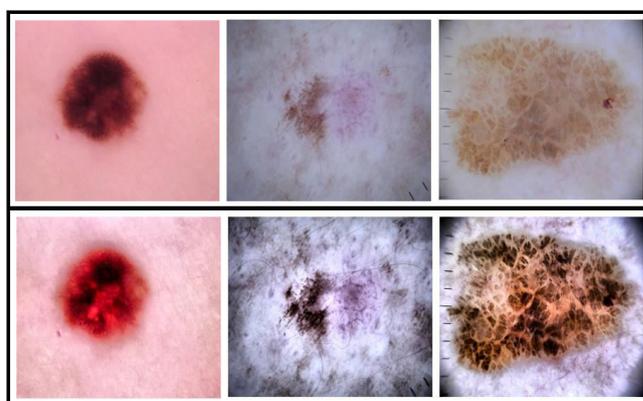


Figure 4. Illustrations before (top) and after (bottom) CLAHE for detection of subtle features.

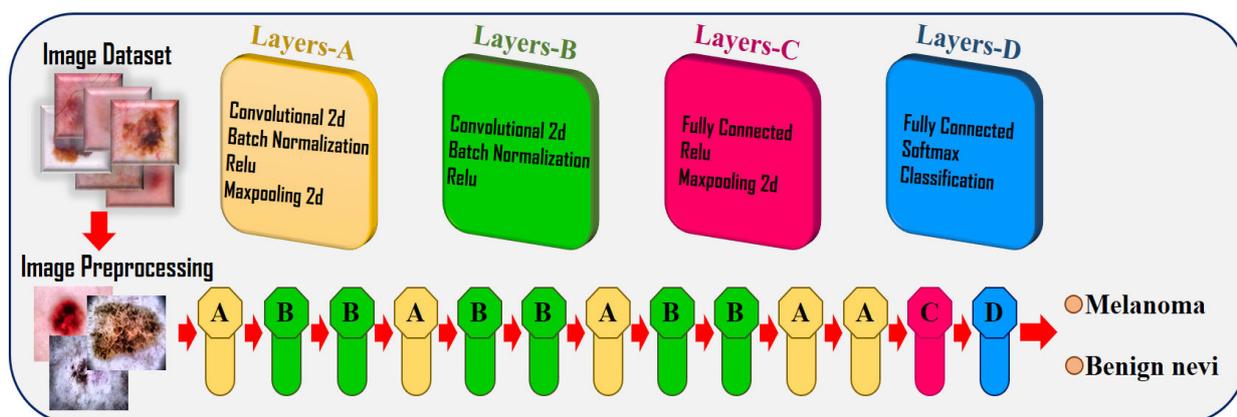


Figure 5. Architecture of the developed CNN model.

The HAM10000 dataset contains skin lesion images with different illumination levels obtained from various patient populations. The illumination level affects the pixel density levels of the images. For this reason, the illumination level can lead to the formation of unnecessary variations that are irrelevant as well as revealing some distinctive features. To increase classifier sensitivity, a contrast-limited adaptive histogram equalization (CLAHE) filtering algorithm was applied via OpenCV to highlight some obscure features of skin lesions in the dataset and to remove some irrelevant features. The results obtained in this preprocessing step are shown in Fig.4. It has been seen with the obtained accuracy results that this digital image preprocessing technique used provides better detection of sensitive point features through convolutional filters that could not be detected by CNN before. It was thought that this preprocessing change made in the dataset could be attributed to the channel-based contrast-enhancing effect of histogram equalization.

### 3.3. DEVELOPED CNN MODEL

The developed CNN model consists of forty five layers in total. Of these, eleven convolution layers have been used for feature maps extraction and twelve Relu layers have been used for activation process. Relu in the activation process includes an activation function that is frequently used in CNN architectures. Six Max-pooling layers have been used to reduce computational cost and reduce data size. It is aimed to reduce the dimensionality by choosing the largest value within the area covered by this filter. In this way, besides the reduced size, unnecessary features are ignored, focusing on retaining more important features. Eleven Batch-normalization layers have been used for the normalization process. This layer provides resistance to the extinction gradient of the convolutional neural network during training. In this way, the training time can be reduced and the model can be provided to produce higher performance. Two Fully-connected layers have been used to start the learning process by taking feature maps as input. A Softmax layer has been used to show the probability that each given input belongs to a class and to produce one-dimensional outputs. Here, the inputs are transformed from a multidimensional matrix format to a single dimension. Finally, a Classification layer has been used to prepare them for the classification process. The architecture of the proposed CNN model is shown in Fig.5.

**Table 1.** Training hiperparameters of CNN model.

Optimization	Epoch	Batch size	Validation freq.	Learning rate
SGD	5	16	8	10 <sup>-4</sup>

The developed model is trained in each of the training processes, which is the epoch, and the most appropriate weights are updated. It has been observed that the training process of the proposed model is fast when the epoch number of the model reaches five. Stochastic Gradient

Descent (SGD) has been chosen as the optimization method which aims to reach the global minimum value by starting with randomly selected variables. By using this optimization method, a higher accuracy value can be obtained, especially in classification processes. Each parameter update in SGD can be calculated as several training samples or as a minibatch compared to a single sample. Since each image of the dataset has a low size, the minibatch sizes are determined as 16. The model has a high accuracy value with increased at a low rate learning level. As the last of the training criteria, the validation frequency value has been determined as 8. In this respect, the visibility of training and loss function curve drawings has been increased. The training values of the model are shown in Table 1. With this model, object recognition and classification problems can be easily solved. In the proposed CNN model, some filters are applied on the images on the convolution layer to obtain the feature maps. The general architectural structure of the CNN model and the parameter values used are given in Table 2.

**Table 2.** Architecture and parameter values of CNN model.

Type	A. Function	Activations	Pooling	Filter size	Stride
Con2d	Relu	55x55x96	Max-P	11, 11	4, 4
Con2d	Relu	27x27x256	-	3, 3	1, 1
Con2d	Relu	27x27x256	-	3, 3	1, 1
Con2d	Relu	27x27x256	Max-P	3, 3	1, 1
Con2d	Relu	14x14x384	-	3, 3	1, 1
Con2d	Relu	14x14x384	-	3, 3	1, 1
Con2d	Relu	14x14x384	Max-P	3, 3	1, 1
Con2d	Relu	7x7x512	-	3, 3	1, 1
Con2d	Relu	7x7x512	-	3, 3	1, 1
Con2d	Relu	7x7x512	Max-P	3, 3	1, 1
Con2d	Relu	2x2x512	Max-P	3, 3	2, 2

An activation function is needed to introduce nonlinear real-world features to CNN. Basically, in a simple CNN,  $x$  is defined as inputs and  $w$  as weights, and activation process  $f(x)$  is applied to the value transferred to the output of the network. This is then considered the final output or the input of another layer. A neural network without an activation function will behave like a linear regression with limited learning capacity. However, the neural network is also required to learn nonlinear situations. For this reason, the Relu activation function linearizes the non-linear outputs from the previous layer and ensures that these outputs are transferred to the next layer as input [30]. Then, the input images have passed through filters on the convolution layers and an activation map has been created. Applying these filters to the images obtained from the previous layer has increased the depth of the network. The size of the filter selected at this stage is important in terms of time complexity because the

smaller the filter size has chosen, the complexity would be greater. In this respect, the processing time will decrease depending on the amount of increasing filter size [31]. However, the large filter size to be selected may also cause a decrease in the desired performance value. Considering this situation, it is of great importance to determine the size of the filter to be applied at this stage. By using Max-pooling layers in the proposed CNN model, it is possible to obtain better feature maps by reducing the computational costs and data size [32]. As yet, very successful results have been obtained by using pre-trained CNN models and many developed models for classification studies in the literature [33, 34]. It is aimed to apply the CNN model developed in this study to other data sets.

### 3.4. PERFORMANCE METRICS AND APPLICATION

The most important thing to do in order for the developed CNN models to be evaluated correctly is that the model is not trained on the entire dataset. In this regard, a typical training/test split is used for this study, with 80% training and 20% testing. Various methods are used to see the performance evaluation of the developed CNN model. As shown in Fig.6, there are four types of outcomes to be obtained with the classification prediction performed. These are true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

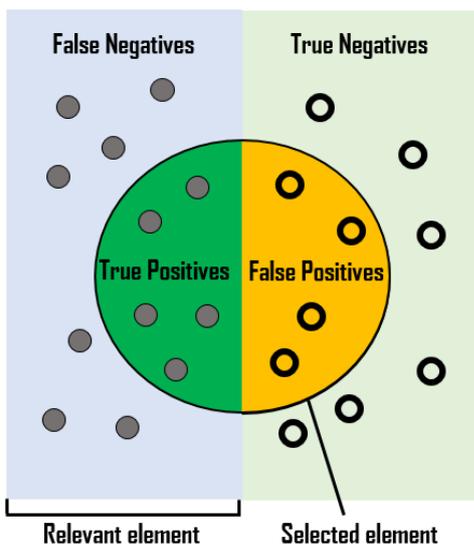


Figure 6. Four main outcomes of classification predictions.

**True Positive (TP):** It is the case that an observation actually belongs to the predicted class. In other words, it is the correct estimation of the melanoma images in the study.

**True Negative (TN):** It is the case that an observation does not actually belong to a class and is predicted not to belong to that class. That is, it is the correct estimation of benign nevi images.

**False Positive (FP):** An observation belonging to the predicted class does not actually belong to that class. In this case, an image belonging to the melanoma class was incorrectly predicted and placed in the benign nevi class.

**False Negative (FN):** It does not belong to a class but is actually assumed to be so. In this case, the image belonging to the benign nevi class was incorrectly predicted and placed in the melanoma class.

Through these metrics, there are commonly used criteria such as Accuracy, Precision, Sensitivity, Specificity, False Positive Rate, False Negative Rate, False Discovery Rate, and  $F_1$ -score value [35]. The formulas of the performance evaluation metrics obtained by using the four basic outcomes are following:

$$Accuracy, (Acc) = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

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

$$Sensitivity, (Sn) = \frac{TP}{TP+FN} \quad (3)$$

$$Specificity, (Sp) = \frac{TN}{TN+FP} \quad (4)$$

$$False\ Positive\ Rate, (FPR) = \frac{FP}{TN+FP} \quad (5)$$

$$False\ Negative\ Rate, (FNR) = \frac{FN}{TP+FN} \quad (6)$$

$$False\ Discovery\ Rate, (FDR) = \frac{FP}{TP+FP} \quad (7)$$

$$F_1 - Score, (F_1) = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (8)$$

These results are obtained over a confusion matrix. This matrix consists of four basic outcome results determined as a result of the estimations made on the test data. After the implemented CNN model is trained, system performance is calculated and the confusion matrix created with other metric functions is revealed [36]. The resulting confusion matrix is given in Fig.7.

Classes	Melanoma	Benign nevi
Melanoma	TP	FP
Benign nevi	FN	TN

Figure 7. Confusion matrix of proposed CNN model.

In this study, a dataset consisting of two classes of images related to skin cancer, melanoma, and benign nevi, was used. The proposed CNN model was trained using 80% of all images of the data set. The proposed model was tested with images of the remaining 20% of the data set. In the process of obtaining the experimental results, a machine with i7-6700K, 4.20 GHz processor, and 16G, 2133MHz RAM was used.

#### IV. EXPERIMENTAL RESULTS AND DISCUSSION

It is aimed to classify these images with the new CNN model developed for the diagnosis of skin cancer through images of skin lesions. Experimental results showed that the proposed CNN model obtained 99.69% accuracy rate and 0.030 loss value as a result of training the data set. Experimental results of accuracy and loss values are shown graphically in Fig.8.

According to the results obtained, the confusion matrix of the CNN model was created in Fig.9. As shown in the Fig.9, a total of 2003 images were used in the testing, and a total of 8012 images were used in the training of the proposed model. Of these training images, 2648 belong to the class of melanoma, and the remaining 5364 to the class of benign nevi. According to the experimental results obtained, 660 of the 662 melanoma images used in the testing phase of the model were placed in the correct class. Only 2 of these images were incorrectly predicted and placed in the class of benign nevi. Likewise, 1338 of the 1341 benign nevi images used for the test phase were predicted correctly, while only 3 were incorrectly predicted. In total, only 5 of the 2003 data used in the testing phase of the data set were predicted incorrectly, while the 1998 images were predicted correctly by the developed model.

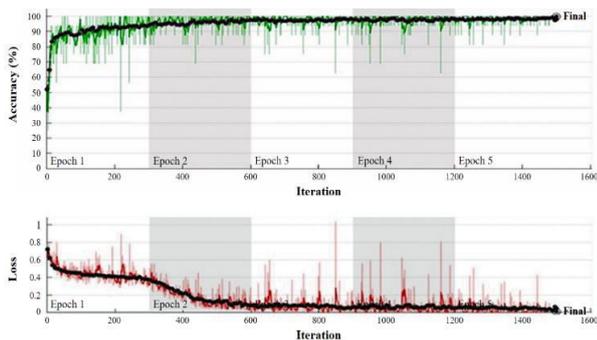


Figure 8. Accuracy and Loss curves of the CNN model.

Classes	Melanoma	Benign nevi
Melanoma	660	2
Benign nevi	3	1338

Figure 9. Confusion matrix results.

The experimental results of the performance metrics calculated using the confusion matrix are given in Table 3.

Table 3. Performance values of CNN model.

Acc	Pre	Sen	Spe	Fpr	Fnr	Fdr	F <sub>1</sub>
99.69	99.70	99.55	99.85	0.00	0.00	0.00	99.62
%	%	%	%	15	45	30	%

Malignant melanoma is one of the most common types of skin cancers today. This type of cancer occurs when precancerous lesions called nevi become malignant melanoma. In this regard, pre-investigation and follow-up of even benign skin lesions for early diagnosis is of great importance in terms of treatment rate and duration of the disease. Today, every individual with skin findings should have a skin cancer test and diagnosis, especially in the early stages. In skin cancers, environmental factors have a great effect besides genetic heritability. It is recommended by experts that cancer patients, patients undergoing treatment, and patients who have completed treatment should have regular testing. However, this process is a troublesome burden for them. Therefore, this heavy workload on experts can be reduced in terms of time and cost with deep learning methods. The recent spread of skin cancers due to environmental effects has necessitated the development of new diagnostic methods. Tests applied using traditional methods are not advantageous in terms of the cost and workload of experts. In addition, diagnostic errors that occur in tests performed using traditional manual methods cause many negativities. Today, diagnosis and classification of skin cancers are the subjects of current studies, since disease data is publicly available and accessible. Studies on the subject in recent years are given in Table 4.

Table 4. Comparison with other study using HAM10000 dataset

Approach	ACC	Method	Year
Ours	99.69%	Developed method	2021
Shahin et al. [37]	89.90%	ResNet-50 + Inception V3	2018
Khan et al. [38]	89.80%	ResNet-50 + ResNet-101	2019
Moldovan [39]	85%	DenseNet121	2019
Mobiny et al. [40]	83.59%	DenseNet-169	2019
Sae-Lim et al. [41]	83.20%	MobileNet	2019
Pai et al. [42]	78%	VGGNet	2019

In the comparative studies, the accuracy results obtained with the well-known training models developed at different times using the same HAM10000 data set were compared. It is thought that comparing the performance of the proposed method with the state-of-the-art methods will be useful in this and future studies. When Table 4 is examined, it can be concluded that the proposed CNN model is successful compared to other methods in terms of reaching 99.69% accuracy. This result showed that the developed model could successfully detect the diagnosis of skin cancer. As in all diseases, early diagnosis plays an important role in the treatment process of skin cancer. In

this respect, this CNN model developed in this field, where computer-aided systems are used effectively for early diagnosis, was able to classify the trained data with high accuracy.

Skin cancer classification approaches using the HAM10000 dataset in recent years have been scanned and summarized in Table 4. In this regard, Shahin et al. achieved 89.90% accuracy on the validation set by combining the outputs of the well-known ResNet50 and InceptionV3 with an ensemble technique [37]. A result close to this approach was obtained by Khan et al. Here, the features in the test dataset are first extracted by a multimodal CNN and then passed through a support vector machine (SVM) with a Radial Basis Function (RBF) kernel [38]. In another study, Moldovan used DenseNet121 as the classification method. The classification method he developed in this approach was composed of a 2-step model divided into three and four classes. An accuracy of 85% in the first and 75% in the second was obtained [39]. Bayesian DenseNet169 was used in an approach by Mobiny et al.. According to the results obtained on the test set, 83.59% accuracy was obtained [40]. The approach taken by Sae-Lim et al. is a modified version of MobileNet. With this method, a lower accuracy result of 83.20% was achieved [41]. In the last approach, Pai et al. created a website application based on VGGNet and achieved an accuracy of 78% for performance evaluation [42].

It can be said that this proposed study achieved higher success than the existing methods in skin lesion classification based on deep learning. The developed classifier offered an effective way to distinguish between melanoma and benign nevi skin lesions with 99.69% accuracy. With 99.69% accuracy, the workload of dermatology experts will be significantly reduced, data loss will be prevented and experts will focus on other issues. This developed CNN model quickly classifies dermatoscopic images and produces high-accuracy results, saving time and cost. Moreover, this developed system can be used for preliminary diagnosis in locations where there is no expert. More work can be done at the second level to differentiate the different types within benign nevi lesions.

## V. CONCLUSION

Early diagnosis of skin cancers is provided by using the developed deep learning model. In this regard, a transfer learning-based framework is presented for the diagnosis of malignant melanoma lesions in the skin. In the developed CNN model, images of melanoma and benign nevi skin lesions were classified using the HAM10000 data set. Deep learning architectures have been widely used in the field of radiology and clinical diagnosis with the developing technology. In this study, a new CNN model is proposed for cancer diagnosis from dermatoscopic skin lesion images. An accuracy value of 99.69% was obtained with the model developed for the detection and classification of cancer images. Thus, it is planned to prevent mistakes made in traditional methods. 8012 images were used for the training phase of the CNN model. The trained model was then tested with 2003

images. Considering that 1998 of the test data was estimated correctly, it was emphasized that the training of the network is of great importance in the developed CNN models. Experimental results show that DenseNet201 achieves about 10% better results than the rest of the networks in measurements, except for hybrid architectures. The instability of the dataset used and the lack of a sufficient number of images are the reason for the lower generalization of DenseNet201 in the classifier. Future work will focus on testing deeper networks and hierarchies in order to distinguish between different types of non-nevi classes.

## REFERENCES

- [1] A. L. Byrd, Y. Belkaid, J. A. Segre, The human skin microbiome. *Nature Reviews Microbiology*, 16(3), 2018, 143-155.
- [2] D. E. O'Sullivan, D. R. Brenner, P. A. Demers, P. J. Villeneuve, C. M. Friedenreich, W. D. King, ComPARE Study Group, Indoor tanning and skin cancer in Canada: A meta-analysis and attributable burden estimation. *Cancer epidemiology*, 59, 2019, 1-7.
- [3] A. J. Miller, M. C. Mihm, Melanoma, *N Engl J Med* 355, 2006, 51-65.
- [4] J. M. Yaiza, R. A. Gloria, G. O. M. Belén, L. R. Elena, J. Gema, M. J. Antonio, B. Houria, Melanoma cancer stem-like cells: Optimization method for culture, enrichment and maintenance. *Tissue and Cell*, 60, 2019, 48-59.
- [5] P. H. Viale, The American Cancer Society's facts & figures: 2020 edition. *Journal of the Advanced Practitioner in Oncology*, 11(2), 2020, 135.
- [6] A. F. Jerant, J. T. Johnson, C. D. Sheridan, T. J. Caffrey, Early detection and treatment of skin cancer. *American family physician*, 62(2), 2000, 357-368.
- [7] K. Narasimhan, V. Elamaran, Wavelet-based energy features for diagnosis of melanoma from dermoscopic images. *International Journal of Biomedical Engineering and Technology*, 20(3), 201, 243-252.
- [8] A. Kulkarni, D. Mukhopadhyay, SVM classifier based melanoma image classification. *Research Journal of Pharmacy and Technology*, 10(12), 2017, 4391-4392.
- [9] V. M. Cohen, E. Pavlidou, J. DaCosta, A. K. Arora, T. Szyszko, M. S. Sagoo, P. Szlosarek, Staging uveal melanoma with whole-body positron-emission tomography/computed tomography and abdominal ultrasound: Low incidence of metastatic disease, high incidence of second primary cancers. *Middle East African journal of ophthalmology*, 25(2), 2018, 91.
- [10] B. Jan, H. Farman, M. Khan, M. Imran, I. U. Islam, A. Ahmad, G. Jeon, Deep learning in big data analytics: a comparative study. *Computers & Electrical Engineering*, 75, 2019, 275-287.
- [11] P. Tschandl, C. Rosendahl, H. Kittler, The HAM10000 dataset, a large collection of multi-

- source dermatoscopic images of common pigmented skin lesions. *Scientific Data*, 5(1), 2018, 1-9.
- [12] F. Nachbar, W. Stolz, T. Merkle, A. B. Cognetta, T. Vogt, M. Landthaler, G. Plewig, The ABCD rule of dermatoscopy: high prospective value in the diagnosis of doubtful melanocytic skin lesions. *Journal of the American Academy of Dermatology*, 30(4), 1994, 551-559.
- [13] E. Harrington, B. Clyne, N. Wesseling, H. Sandhu, L. Armstrong, H. Bennett, T. Fahey, (2017). Diagnosing malignant melanoma in ambulatory care: a systematic review of clinical prediction rules. *BMJ open*, 7(3), 2017, e014096.
- [14] M. H. Jafari, N. Karimi, E. Nasr-Esfahani, S. Samavi, S. M. R. Soroushmehr, K. Ward, K. Najarian, Skin lesion segmentation in clinical images using deep learning. In *2016 23rd International conference on pattern recognition (ICPR)*, 337-342.
- [15] V. Yadav, V. D. Kaushik, Detection of melanoma skin disease by extracting high level features for skin lesions. *International Journal of Advanced Intelligence Paradigms*, 11(3-4), 2018, 397-408.
- [16] M. Ruela, C. Barata, J. S. Marques, J. Rozeira, A system for the detection of melanomas in dermoscopy images using shape and symmetry features. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 5(2), 2017, 127-137.
- [17] M.P.P. Meena, T. Balaji, Advanced Method Using Find and Detection Skin Cancer Classification. *International Journal of Advanced Networking & Applications (IJANA)*, 08(05), 2017, 24-27.
- [18] K. Thurnhofer-Hemsi, Domínguez, Analyzing digital image by deep learning for melanoma diagnosis. In *2019 International Work-Conference on Artificial Neural Networks*, 270-279.
- [19] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, S. Thrun, Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 2017, 115-118.
- [20] T. Zhou, K. H. Thung, X. Zhu, D. Shen, Effective feature learning and fusion of multimodality data using stage-wise deep neural network for dementia diagnosis. *Human brain mapping*, 40(3), 2019, 1001-1016.
- [21] L. Zhang, F. F. Yin, J. Cai, A multi-source adaptive MR image fusion technique for MR-Guided radiation therapy. *International Journal of Radiation Oncology, Biology, Physics*, 102(3), 2018, e552.
- [22] M. El Adoui, S. A. Mahmoudi, M. A. Larhman, M. Benjelloun, MRI breast tumor segmentation using different encoder and decoder CNN architectures. *Computers*, 8(3), 2019, 52.
- [23] R. Rouhi, M. Jafari, S. Kasaei, P. Keshavarzian, Benign and malignant breast tumors classification based on region growing and CNN segmentation. *Expert Systems with Applications*, 42(3), 2015, 990-1002.
- [24] R. Guerrero, C. Qin, O. Oktay, C. Bowles, L. Chen, R. Joules, D. Rueckert, White matter hyperintensity and stroke lesion segmentation and differentiation using convolutional neural networks. *NeuroImage: Clinical*, 17, 2018, 918-934.
- [25] Z. Gao, X. Wang, S. Sun, D. Wu, J. Bai, Y. Yin, V. H. C. de Albuquerque, Learning physical properties in complex visual scenes: an intelligent machine for perceiving blood flow dynamics from static CT angiography imaging. *Neural Networks*, 123, 2020, 82-93.
- [26] Z. Gao, J. Chung, M. Abdelrazek, S. Leung, W. K. Hau, Z. Xian, S. Li, S. Privileged modality distillation for vessel border detection in intracoronary imaging. *IEEE transactions on medical imaging*, 39(5), 2019, 1524-1534.
- [27] M. Prathiba, D. Jose, R. Saranya, Automated melanoma recognition in dermoscopy images via very deep residual networks. In *2019 IOP Conference Series: Materials Science and Engineering*.
- [28] M. Abd El Aziz, A. A. Ewees, A. E. Hassanien, Whale optimization algorithm and moth-flame optimization for multilevel thresholding image segmentation. *Expert Systems with Applications*, 83, 2017, 242-256.
- [29] S. Mirjalili, A. Lewis, The whale optimization algorithm. *Advances in engineering software*, 95, 2016, 51-67.
- [30] Y. D. Zhang, C. Pan, J. Sun, C. Tang, Multiple sclerosis identification by convolutional neural network with dropout and parametric ReLU. *Journal of computational science*, 28, 2018, 1-10.
- [31] E. Limonova, A. Sheshkus, A. Ivanova, D. Nikolaev, Convolutional neural network structure transformations for complexity reduction and speed improvement. *Pattern Recognition and Image Analysis*, 28(1), 2018, 24-33.
- [32] M. V. Valueva, N. N. Nagornov, P. A. Lyakhov, G. V. Valuev, N. I. Chervyakov, N. I. Application of the residue number system to reduce hardware costs of the convolutional neural network implementation. *Mathematics and Computers in Simulation*, 177, 2020, 232-243.
- [33] K. Thurnhofer-Hemsi, E. Dominguez, A convolutional neural network framework for accurate skin cancer detection. *Neural Processing Letters*, 2020, 1-21.
- [34] N. Zhang, Y. X. Cai, Y. Y. Wang, Y. T. Tian, X. L. Wang, B. Badami, Skin cancer diagnosis based on optimized convolutional neural network. *Artificial intelligence in medicine*, 102, 2020, 101756.
- [35] D. Jain, V. Singh, Feature selection and classification systems for chronic disease prediction: A review. *Egyptian Informatics Journal*, 19(3), 2018, 179-189.
- [36] A. Luque, A. Carrasco, A. Martín, A. de las Heras, The impact of class imbalance in classification performance metrics based on the binary confusion matrix. *Pattern Recognition*, 91, 2019, 216-231.

- [37] A. H. Shahin, A. Kamal, M. A. Elattar, Deep ensemble learning for skin lesion classification from dermoscopic images. *In 2018 9th Cairo International Biomedical Engineering Conference (CIBEC)*, 2018, 150-153.
- [38] M. A. Khan, M. Y. Javed, M. Sharif, T. Saba, A. Rehman, Multi-model deep neural network based features extraction and optimal selection approach for skin lesion classification, *In 2019 international conference on computer and information sciences (ICCIS)*, 2019, 1-7.
- [39] D. Moldovan, Transfer learning based method for two-step skin cancer images classification. *In 2019 E-Health and Bioengineering Conference (EHB)*, 2019, 1-4.
- [40] A. Mobiny, A. Singh, H. Van Nguyen, Risk-aware machine learning classifier for skin lesion diagnosis. *Journal of clinical medicine*, 8(8), 2019, 1241.
- [41] W. Sae-Lim, W. Wettayaprasit, P. Aiyarak, Convolutional neural networks using mobilenet for skin lesion classification, *In 2019 16th international joint conference on computer science and software engineering (JCSSE)2019*, 242-247, IEEE.
- [42] K. Pai, A. Giridharan, Convolutional Neural Networks for classifying skin lesions, *In TENCON 2019-2019 IEEE Region 10 Conference 2019*, 1794-1796.

#### AUTHOR'S BIOGRAPHY

##### **Erdal Özbay**



He received the bachelor's degree in computer engineering from American University, TRNC, in 2010. He received the M.S. and Ph.D. degree in Computer Engineering from Firat University, Turkey, in 2013 and 2018 respectively. He is currently an Assistant Professor with the

Department of Computer Engineering, Firat University. His research interests include deep learning, image processing, and optimization techniques.

##### **F. Altunbey Özbay**



She received the bachelor's degree in computer engineering from Firat University, Turkey, in 2013. She received the M.S. and Ph.D. degree in Software Engineering from Firat University, Turkey, in 2015 and 2020 respectively. She is currently an

Assistant Professor with the Department of Software Engineering, Firat University. Her research interests include artificial neural networks, data mining, and metaheuristic optimization.