



RESEARCH PAPER

Artificial intelligence-assisted detection model for melanoma diagnosis using deep learning techniques

Hediye Orhan ^{1, ‡} and Emrehan Yavşan ^{2, *, ‡}

¹Department of Computer Engineering, Necmettin Erbakan University, 42140 Konya, Türkiye,

²Department of Electronics and Automation, Tekirdağ Namık Kemal University, Tekirdağ, Türkiye

*Corresponding Author

‡hediyeorhan2015@gmail.com (Hediye Orhan); eyavsan@nku.edu.tr (Emrehan Yavşan)

Abstract

The progressive depletion of the ozone layer poses a significant threat to both human health and the environment. Prolonged exposure to ultraviolet radiation increases the risk of developing skin cancer, particularly melanoma. Early diagnosis and vigilant monitoring play a crucial role in the successful treatment of melanoma. Effective diagnostic strategies need to be implemented to curb the rising incidence of this disease worldwide. In this work, we propose an artificial intelligence-based detection model that employs deep learning techniques to accurately monitor nevi with characteristics that may indicate the presence of melanoma. A comprehensive dataset comprising 8598 images was utilized for the model development. The dataset underwent training, validation, and testing processes, employing the algorithms such as AlexNet, MobileNet, ResNet, VGG16, and VGG19, as documented in current literature. Among these algorithms, the MobileNet model demonstrated superior performance, achieving an accuracy of 84.94% after completing the training and testing phases. Future plans involve integrating this model with a desktop program compatible with various operating systems, thereby establishing a practical detection system. The proposed model has the potential to aid qualified healthcare professionals in the diagnosis of melanoma. Furthermore, we envision the development of a mobile application to facilitate melanoma detection in home environments, providing added convenience and accessibility.

Keywords: Artificial intelligence; deep learning; machine learning; melanoma detection; skin cancer

AMS 2020 Classification: 68T01; 68T07; 68T40; 92B20

1 Introduction

Skin cancer is a disease that can have serious consequences if left untreated. Especially an aggressive type of skin cancer, such as melanoma, can spread quickly and metastasize to other organs. Skin cancer can be fatal if not diagnosed in time and treated appropriately. Because

melanoma is more aggressive than other types of skin cancer and tends to metastasize, treatment becomes more difficult when diagnosed in advanced stages. Even in cases of melanoma treated with early diagnosis, in some cases, the disease can progress and become life-threatening. Because skin cancer can have serious consequences, it is important to take preventive measures by reducing risk factors such as regular skin examinations and sun exposure. Early detection can make treatment more effective and prevent serious consequences. For the early diagnosis of this disease, image processing [1, 2] and deep learning methods can be applied. Many studies have reported the application of artificial intelligence (AI) algorithms in the detection of various types of cancer [3–5], including skin cancer [6, 7].

Yildiz [8], proposes an automatic detection system for melanoma using deep learning methods, specifically model C4Net. Sultana and Puhan [9] review the use of deep learning techniques to detect melanoma from other skin lesions in clinical and dermoscopy images. They emphasized that deep learning techniques outperform traditional methods in skin cancer detection, but data labeling for deep learning techniques is challenging. A similar study to Sultana and Puhan was conducted by Poorna et al. Poorna et al. [10] discuss the development of a computer vision-aided system for the early diagnosis of melanoma, a type of skin cancer. The study compares the accuracy and precision of conventional supervised learning techniques with deep learning-based methods for melanoma detection. Their learning techniques used for the classification were Total Dermoscopic Score, K Nearest Neighbor (KNN) [11–13] and, Support Vector Machine (SVM) [14]. Kwiatkowska et al. [15] detected melanoma from dermoscopic images using ResNet and its different versions. However, higher accuracy rates can be achieved with different models. Shchetinin et al. [16] developed a computer-aided detection system for Melanoma detection with Deep Neural Networks on HAM10000 dataset. By combining multiple dataset instead of a single dataset, a system more suitable for real scenarios can be presented.

In this study, we developed a system for the early detection of melanoma using various deep-learning techniques. The system is tested on a vast dataset. This dataset is a combination of 2 different datasets to increase the reality of the system and its feasibility for clinical applications. Training and testing were performed on a relatively rich dataset in terms of the variety and number of datasets. The presented system will play an active role in the early detection of skin cancer, which is likely to threaten our lives even more in the future, and will contribute to the development of future research on the diagnosis of this disease.

2 General properties of method

Preparing the dataset

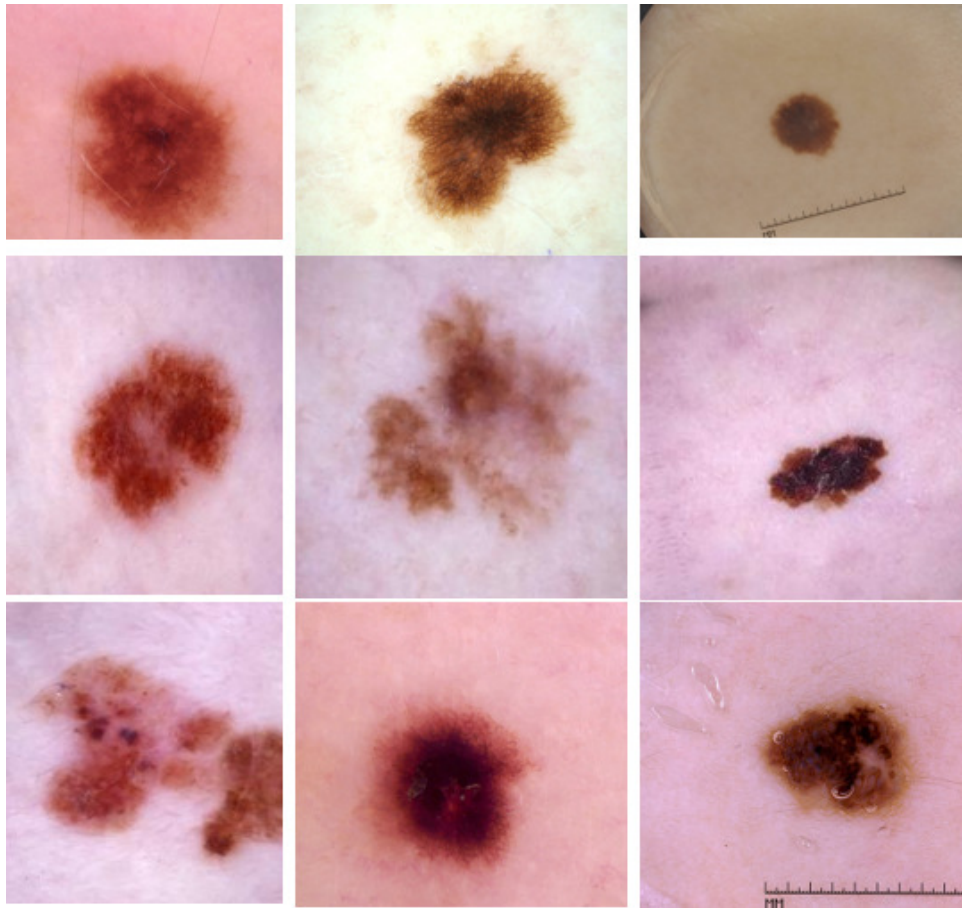
In this study, two different datasets were blended. In this way, a rich dataset was obtained in terms of data diversity. The nevus category data from the first dataset called nevus classifier [17] was added to the new dataset for healthy nevus images. In the second dataset, called skin lesions [18], the data in the melanoma category were added to the same new dataset as melanoma images for cancerous nevus. Thus, a dataset containing healthy, labeled as nevus, and cancerous, labeled as melanoma, nevus images were obtained.

The dataset consists of two classes: nevus and melanoma. The melanoma class represents cancerous mole images and the nevus class represents healthy mole images. There are a total of 8598 data in the dataset. The distribution of the dataset is shown in Table 1. Each class in the dataset is divided into 60% train, 20% test, and 20% validation.

Figure 1 illustrates some of the images in the dataset. In this work, since the dataset is rich and diverse enough, there was no need to augment the data.

Table 1. The data distribution in the dataset

TRAIN	
Melanoma	Nevus
2328	3174
VALIDATION	
Melanoma	Nevus
582	794
TEST	
Melanoma	Nevus
728	992

**Figure 1.** Some of the dataset images

Proposed artificial intelligence models

The proposed AI model for the presented problem is based on a Convolutional Neural Network (CNN). CNNs [19–21] are a deep learning discipline that has proven its success in computer vision and has models designed for various problems. They can be used in vision systems of robots and autonomous vehicles for face [22, 23], object [24] and, traffic sign [25, 26] recognition. A CNN generally consists of Convolution, Pooling, and Fully Connected Layer structures. In a CNN, the image is directly input into the network, followed by several convolution and pooling processes. The outputs of the convolution and pooling processes feed one or more fully connected layers. Eventually, the class label is extracted as an output. Figure 2 depicts a general CNN structure for the detection problem in this paper. The convolutional layers shown in the CNN diagram in Figure 2 serve as feature extractors. Thus, the features of the input images are extracted. Neurons

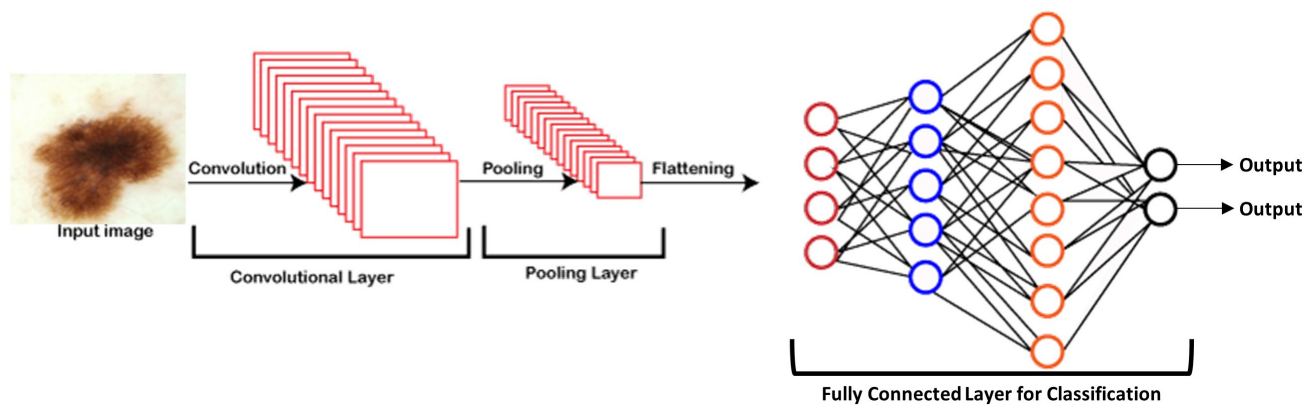


Figure 2. Customized CNN schematic for the classification problem

in the convolutional layers are organized into feature maps. The aim of the pooling layer is to reduce the size (width x height) of the input image. This simplifies the computation for the next layer and also prevents overfitting. Filters similar to those used in the previous layer, convolution, are also used in the pooling layer. In these filters, the image matrix is shifted and the highest value of the pixels (maximum pooling) or the average (average pooling) is calculated. Maximum pooling is often used because it produces higher performance due to the simplicity and rapidity of the process.

The fully connected layer is the last layer of the CNN structure. In the CNN structure, the fully connected layer follows the convolution and pooling layers in succession. A fully connected layer is a standard layer used in classification problems [27]. Convolution and pooling layers can be used many times before reaching the fully connected layer. However, the matrices at the output of the convolution and pooling layers need to be flattened in order to be used in the fully connected layer. Therefore, a flattening layer must be used before the fully connected layer. In the flattening layer, the matrices from the convolution and pooling layers are converted into a vector. This process is called flattening.

For the classification problem in this study; AlexNet, MobileNet, ResNet, VGG16 [28, 29] and VGG19 models were used. These models were adapted to the presented problem by making the revisions indicated in Figure 3. The nevus images obtained from Kaggle were sized 224 x 224 and were input to the models. For each model, sigmoid was chosen as the activation function because it is more suitable for two-class problems. Figure 3 symbolizes the revisions made to the models proposed. In future studies, the softmax activation function may be preferred for classification problems with a larger number of outputs, as in the 1000-class model in Figure 3.

3 Training and performance evaluation of models

The training of the CNN-based models using the sigmoid activation function was performed for various values of the epoch, batch size, and learning rate parameters. For each model, the weights that provided the highest accuracy value on the validation data at the end of the training were determined. The testing process was carried out with these weights. Thus, test performance was improved and over-fitting was prevented. The algorithm memorizes the data in the training set, which is over-fitting. The algorithm that memorizes the training data cannot perform well enough on the test data. Therefore, the determination of the training parameters is crucial. Table 2 lists the models proposed for this work and the parameters used in the training of these models. The learning rate parameters are kept low to prevent over-fitting. Epoch values are set between 10-30. Also, an early stop was included in the algorithm during the training of some of the models. When the training is terminated early (early stop), the training accuracy must be more than 95%.

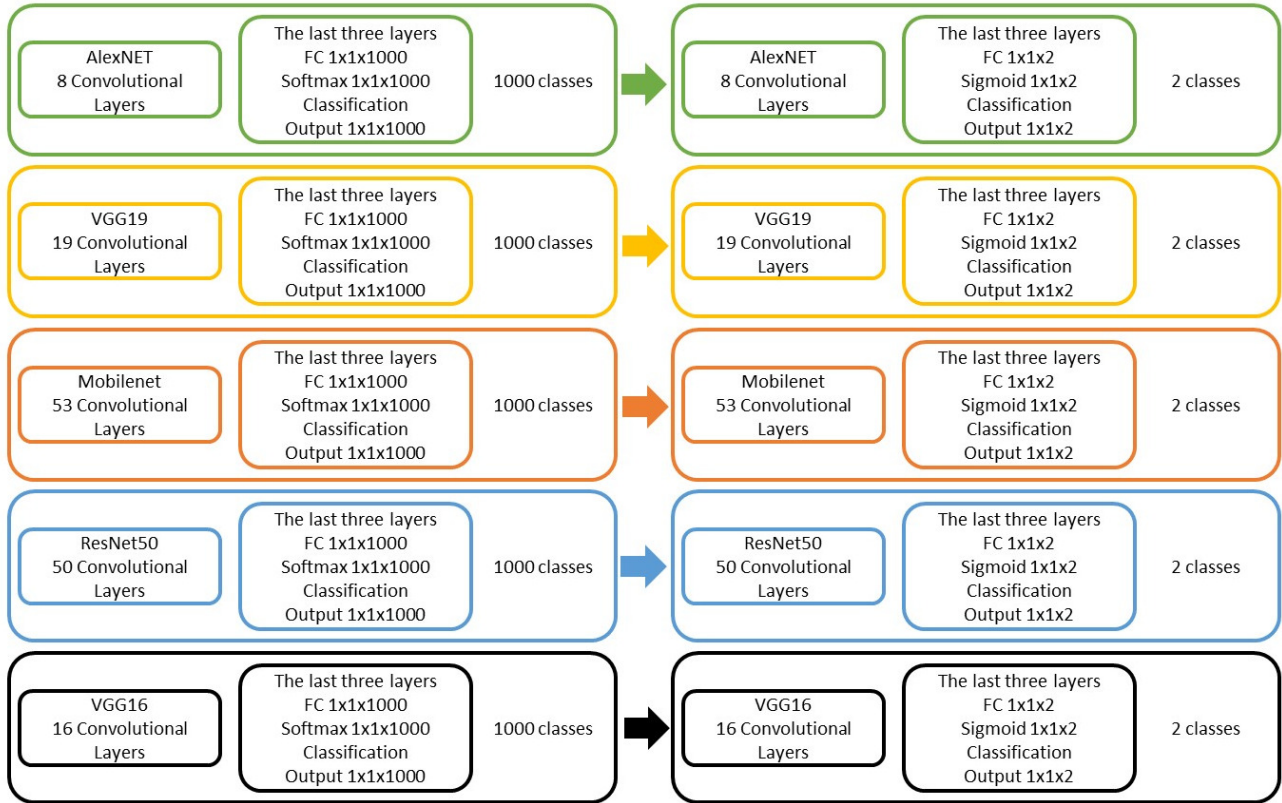


Figure 3. Revised CNN models

The parameters that produce the maximum accuracy on the validation data are assigned after the termination of the training. This is one strategy used to avoid over-fitting.

$$precision = \frac{TP}{TP + FP'} \quad (1)$$

$$recall = \frac{TP}{TP + FN'} \quad (2)$$

$$accuracy = \frac{TP + TN}{TP + FN + TN + FP'} \quad (3)$$

$$specificity = \frac{TN}{TN + FP'} \quad (4)$$

$$F1_{score} = 2 \cdot \frac{precision \cdot recall}{precision + recall} \quad (5)$$

The performance of the proposed models was evaluated through the metrics in Eqs (1) - (5). The

symbols in the metrics can be expressed as 1 and 0 as follows. True Positive (TP) means that the true value is 1 and the predicted value is 1, True Negative (TN) means that the true value is 0 and the predicted value is 0, False Positive (FP) means that the true value is 0 but the predicted value is 1 and False Negative (FN) means that the true value is 1 but the predicted value is 0. The performances of the models are reported in the same list next to the hyper-parameters in Table 2.

Table 2. Hyper-parameters used to train the models and performance metrics presented to evaluate the test performance of the models

Proposed Models	Hyper-parameters for Training			Performance Metrics				
	Epoch	Batch Size	Learning Rate	Accuracy	Precision	Recall	Specificity	F1 Score
AlexNet	30	16	0.00015	0.6767	0.8263	0.5564	0.8406	0.665
MobileNet	10	16	0.000002	0.8494	0.7929	1	0.6442	0.8845
MobileNet	10	8	0.000002	0.8494	0.7929	1	0.6442	0.8845
ResNet	30	16	0.00015	0.8110	0.7683	0.96270	0.6043	0.8545
ResNet	60	16	0.00015	0.8203	0.7729	0.9747	0.6098	0.8622
VGG16	30	16	0.00015	0.8337	0.7936	0.9616	0.6593	0.8696
VGG19	30	16	0.00015	0.8279	0.7979	0.9395	0.6758	0.8629

In neural networks [30, 31], the training parameters are used for high performance, which can be set similarly to the hyper-parameters in Table 2. The absolute minimum value of the error needs to be found. In this process, various optimizers [32–35] are employed to reduce the computational load and minimize losses in the training process. In the present work, Adaptive Moment Estimation (Adam) Optimizer [33] is preferred. Adam algorithm is not only fast but also has low memory usage [36].

MobileNet achieved the highest performance with close to 85% accuracy as a result of the tests of the models on the GPU. Afterwards, the batch size value for the MobileNet model was increased to save time. Figure 4 shows the ROC curve of the model. The area under the curve (AUC) can be considered as an indicator of the model's performance. The higher the AUC, the better the performance of the trained model.

Besides the metrics expressed by Eqs. (1) - (5), the confusion matrix in Table 3 is utilized to evaluate the test performance of the MobileNet model. The confusion matrix is composed of TP, TN, FP and FN values calculated for the testing process.

Table 3. Test performance of the proposed model based on confusion matrix

		Model Prediction	
		Melanoma	Nevus
Actual Output	Melanoma	469 (TN)	259 (FP)
	Nevus	0 (FN)	992 (TP)

The test performance of the proposed model is evaluated on 728 cancerous and 992 healthy nevus images that are not in the training dataset. While the model detected 259 out of 728 melanoma images as healthy, was able to detect all 992 healthy mole images successfully (without failure).

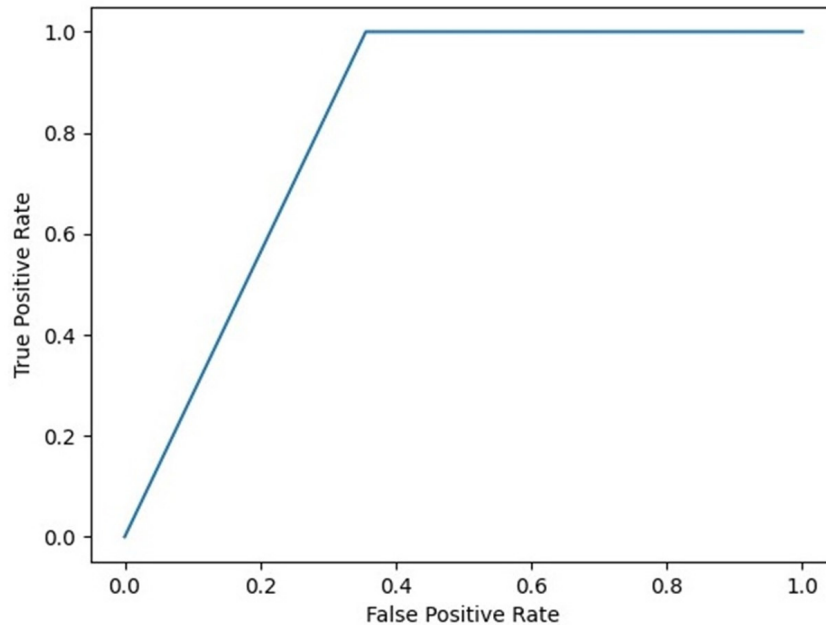


Figure 4. ROC curve of the MobileNet model

4 Results and discussion

In this work, an AI-supported detection model has been developed for skin cancer, the incidence of which is increasing worldwide. AlexNet, MobileNet, ResNet, VGG16 and VGG19 models are employed for the introduced classification problem for the early detection of skin cancer. The proposed models have been trained and tested with a rich data set. The performance of the models has been evaluated based on performance metrics frequently used in the literature. The MobileNet model has offered the highest performance out of these models. The model has achieved 84.94% accuracy after testing.

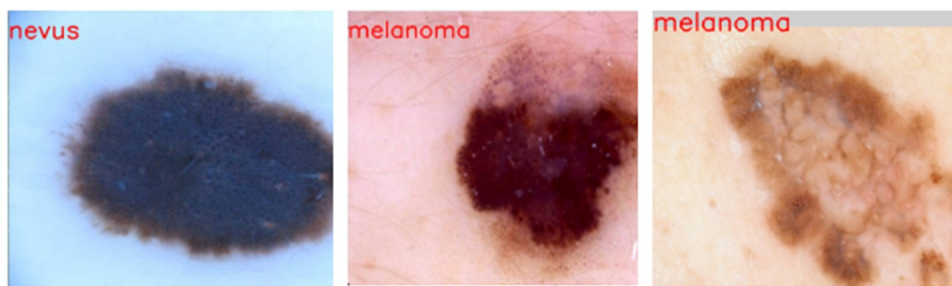


Figure 5. Classification outputs of the model

Table 4 has been prepared to compare the performance of the model with benchmark works in the literature. Successful classifications have been performed with the model file obtained at the end of the training. Also, some of the classified images have been illustrated in Figure 5.

Early diagnosis of skin cancer, aided by these and similar studies, will prevent deaths from this disease. This paper proves that deep learning algorithms can be successfully used in the detection of skin cancer. As the number and diversity of the dataset increase, the success and accuracy of the deep learning model increase. However, labeling the data becomes more difficult. Also, the proposed deep learning models should be considered as an assistant. The use of these models should not be independent of experts, especially in the field of health.

Table 4. Comparison of the model proposed in this study with competitive alternatives in the literature

Studies (Year)	Dataset (Image Counts)	Algorithm	Accuracy
Sultana and Puhan (2018)	ISIC	Deep learning+Sparse Coding +SVM	74.30%
		Deep learning ensemble	80.70%
	Dermofit Image Library	Multi-resolution-tract CNN	79.50%
		Multi-scale feature extraction using full-CNN	81.80%
	MED-NODE (170)	5 layered CNN	81.00%
	MoleMap	ResNet+bilinear pooling	71.00%
	ISBI 2016 (900)	Fine-tuning using VGGNet	81.30%
		CNN+Fisher Encoding	83.09%
Poorna et al. (2020)	ISIC (1600)	TDS	57.30%
		KNN	66.90%
		SVM	66.35%
Kwiatkowska et al. (2021)	ISIC 2018	ResNet	72.57%
		ResNetXt	77.09%
		SE-ResNet	81.62%
		SE-ResNetXt	0.721 ±0.007
Shchetinin et al. (2021)	HAM10000	InceptionV3	72.20%
		MobileNet	78.20%
Proposed study	Kaggle Datasets (8598)	MobileNet	84.94%

Declarations

Ethical approval

The authors state that this research complies with ethical standards. This research does not involve either human participants or animals.

Consent for publication

Not applicable.

Conflicts of interest

The authors declare that they have no conflict of interest.

Data availability statement

Data availability is not applicable to this article as no new data were created or analysed in this study.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author's contributions

H.O.: Methodology, Software, Conceptualization, Validation, Investigation, Writing - Original Draft, Visualization, Data Curation. E.Y.: Methodology, Conceptualization, Validation, Investigation, Writing - Original Draft, Writing - Review & Editing, Visualization, Supervision, Project Administration. All authors discussed the results and contributed to the final manuscript.

Acknowledgements

Not applicable.

References

- [1] Jain, S. and Pise, N. Computer aided melanoma skin cancer detection using image processing, *Procedia Computer Science*, 48, 735-740, (2015). [[CrossRef](#)]
- [2] Kaymak, R., Kaymak, C. and Ucar, A. Skin lesion segmentation using fully convolutional networks: A comparative experimental study. *Expert Systems with Applications*, 161, 113742, (2020). [[CrossRef](#)]
- [3] Tabrizchi, H., Parvizpour, S. and Razmara, J. An improved VGG model for skin cancer detection. *Neural Processing Letters*, 1-18, (2022). [[CrossRef](#)]
- [4] Kılıç, A.E. and Karakoyun, M. Breast cancer detection using machine learning algorithms. *International Journal of Advanced Natural Sciences and Engineering Researches*, 7(3), 91-95, (2023). [[CrossRef](#)]
- [5] Dildar, M., Akram, S., Irfan, M., Khan, H.U., Ramzan, M., Mahmood, A.R. et al. Skin cancer detection: a review using deep learning techniques. *International Journal of Environmental Research and Public Health*, 18(10), 5479, (2021). [[CrossRef](#)]
- [6] Skin Cancer Foundation. Melanoma. <https://www.skincancer.org/skin-cancer-information/melanoma/>, 2021, Access Date: 13th April 2023.
- [7] American Cancer Society. What Causes Melanoma Skin Cancer? <https://www.cancer.org/cancer/melanoma-skin-cancer/causes-risks-prevention/what-causes.html>, 2022, Access Date: 25th April 2023.
- [8] Yildiz, O. Melanoma detection from dermoscopy images with deep learning methods: A comprehensive study. *Journal of the Faculty of Engineering and Architecture of Gazi University*, 34(4), 1-42, (2019). [[CrossRef](#)]
- [9] Sultana, N.N., Puhan, N.B. Recent deep learning methods for melanoma detection: a review. In *Communications in Computer and Information Science (Mathematics and Computing)* (vol. 834) pp. 118-132, Singapore: Springer, (2018). [[CrossRef](#)]
- [10] Poorna, S.S., Reddy, M.R.K., Akhil, N., Kamath, S., Mohan, L., Anuraj, K. and Pradeep, H.S. Computer vision aided study for melanoma detection: a deep learning versus conventional supervised learning approach. In *Proceedings, Advanced Computing and Intelligent Engineering: Proceedings of ICACIE 2018* (Vol. 1) pp. 75-83, Singapore: Springer, (2020). [[CrossRef](#)]
- [11] Yavşan, E. and Ucar, A. Teaching human gestures to humanoid robots by using Kinect sensor. In *Proceedings, 2015 23rd Signal Processing and Communications Applications Conference (SIU)*, pp. 1208-1211, Malatya, Türkiye, (2015, May). [[CrossRef](#)]
- [12] Yavşan, E. and Ucar, A. Gesture imitation and recognition using Kinect sensor and extreme learning machines. *Measurement*, 94, 852-861, (2016). [[CrossRef](#)]
- [13] Baitu, G.P., Gadalla, O.A.A. and Öztekin, Y.B. Traditional machine learning-based classification of cashew kernels using colour features. *Journal of Tekirdag Agricultural Faculty*, 20(1), 115-124, (2023). [[CrossRef](#)]
- [14] Ucar, A. and Özalp, R. Efficient android electronic nose design for recognition and perception of fruit odors using Kernel Extreme Learning Machines. *Chemometrics and Intelligent Laboratory Systems*, 166, 69-80, (2017). [[CrossRef](#)]
- [15] Kwiatkowska, D., Kluska, P. and Reich, A. Convolutional neural networks for the detection of

malignant melanoma in dermoscopy images. *Advances in Dermatology and Allergology/Postepy Dermatol Alergol*, 38(3), 412, (2021). [[CrossRef](#)]

- [16] Shchetinin, E.Y., Sevastianov, L.A., Kulyabov, D.S., Ayryan, E.A. and Demidova, A.V. Melanoma detection computer system development with deep neural networks. In *Distributed Computer and Communication Networks: Control, Computation, Communications (Mathematics and Computing)* (vol. 1337) pp. 422-434, Singapore: Springer, (2020). [[CrossRef](#)]
- [17] <https://www.kaggle.com/datasets/mathewmarcum/nevousclassifier>, 2020, Access Date: 22th February 2023.
- [18] <https://www.kaggle.com/datasets/wanderdust/skin-lesion-analysis-toward-melanoma-detection>, 2020, Access Date: 20th February 2023.
- [19] Altekin, F. and Demir, H. Emotion detection from facial expression using different feature descriptor methods with Convolutional Neural Networks. *European Journal of Engineering and Applied Sciences*, 4(1), 14-17, (2021). [[CrossRef](#)]
- [20] Efe, E. and Ozsen, S. CoSleepNet: Automated sleep staging using a hybrid CNN-LSTM network on imbalanced EEG-EOG datasets. *Biomedical Signal Processing and Control*, 80, 104299, (2023). [[CrossRef](#)]
- [21] Tajbakhsh, N., Roth, H., Terzopoulos, D. and Liang, J. Guest editorial annotation-efficient deep learning: the holy grail of medical imaging. *IEEE transactions on medical imaging*, 40(10), 2526-2533, (2021). [[CrossRef](#)]
- [22] Tatar, A.B. Biometric identification system using EEG signals. *Neural Computing and Applications*, 35(1), 1009-1023, (2023). [[CrossRef](#)]
- [23] Ucar, A. Deep Convolutional neural networks for facial expression recognition. In *Proceedings, 2017 IEEE International Conference on Innovations in Intelligent Systems and Applications (INISTA)*, pp. 371-375, Gdynia, Poland, (2017, July). [[CrossRef](#)]
- [24] Aslan, S.N., Özalp, R., Uçar, A. and Güzeliş, C. New CNN and hybrid CNN-LSTM models for learning object manipulation of humanoid robots from demonstration. *Cluster Computing*, 25(3), 1575-1590, (2022). [[CrossRef](#)]
- [25] Küçük, Ö., Gökçe, B. and Yavşan, E. Otonom tabanlı işaret ve şerit tanımak amacı ile bir öğrenme sisteminin geliştirilmesi. *International Journal of Engineering Research and Development*, 13(3), 19-25, (2021). [[CrossRef](#)]
- [26] Uçar, A., Demir, Y. and Güzeliş, C. Object recognition and detection with deep learning for autonomous driving applications. *Simulation*, 93(9), 759-769, (2017). [[CrossRef](#)]
- [27] Basha, S.H.S., Dubey, S.R., Pulabaigari, V. and Mukherjee, S. Impact of fully connected layers on performance of convolutional neural networks for image classification. *Neurocomputing*, 378, 112-119, (2020). [[CrossRef](#)]
- [28] Polat, D.S. and Yavşan, E. Geri dönüşümdeki verimliliği artırmak için yapay zeka destekli atık tespit sistemi. In *Proceedings, 1st International Conference on Engineering, Natural and Social Sciences (ICENSOS 2022)*, pp. 757-761, Konya, Türkiye, (2022, December).
- [29] Orhan, H. and Yavşan, E. A machine learning assisted monitoring system for early detection of melanoma. In *Proceedings, 7th International Conference on Computational Mathematics and Engineering Sciences*, pp. 110, Elazığ, Türkiye, (2023, May).
- [30] Öztürk, P., Alisoy, H. and Mutlu, R. Prediction of CAT 6A U/FTP cable parameters produced by double twist and triple twist machines using artificial neural networks, and comparison of the predicted results. *European Journal of Engineering and Applied Sciences*, 2(2), 41-51, (2019).

- [31] Kara, M.R. and Yavşan, E. Kapasitif bir algılayıcı üzerinden demodülasyon hatalarının yapay sinir ağları kullanılarak düşürülmesi. In Proceedings, *1st International Conference on Engineering and Applied Natural Sciences (ICEANS'22)*, pp. 1705-1708, Konya, Türkiye, (2022, May).
- [32] Reddy, R.V.K., Rao, B.S. and Raju, K.P. Handwritten Hindi digits recognition using convolutional neural network with RMSprop optimization. In Proceedings, *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 45-51, Madurai, India, (2018, June). [[CrossRef](#)]
- [33] Bock, S. and Weiß, M. A proof of local convergence for the Adam optimizer. In Proceedings, *2019 International Joint Conference on Neural Networks (IJCNN)*, pp. 1-8, Budapest, Hungary, (2019, July). [[CrossRef](#)]
- [34] Lydia, A. and Francis, S. Adagrad—an optimizer for stochastic gradient descent. *International Journal of Information and Computing Science*, 6(5), 566-568, (2018). [[CrossRef](#)]
- [35] Wang, Y., Liu, J., Mišić, J., Mišić, V.B., Shaohua, L. and Chang, X. Assessing optimizer impact on DNN model sensitivity to adversarial examples. *IEEE Access*, 7, 152766-152776, (2019). [[CrossRef](#)]
- [36] Tato, A., Nkambou, R. Improving Adam optimizer. In Proceedings, *ICLR 2018 Workshop*, (2018). [[CrossRef](#)]

Mathematical Modelling and Numerical Simulation with Applications (MMNSA)

(<https://dergipark.org.tr/en/pub/mmnsa>)



Copyright: © 2023 by the authors. This work is licensed under a Creative Commons Attribution 4.0 (CC BY) International License. The authors retain ownership of the copyright for their article, but they allow anyone to download, reuse, reprint, modify, distribute, and/or copy articles in MMNSA, so long as the original authors and source are credited. To see the complete license contents, please visit (<http://creativecommons.org/licenses/by/4.0/>).

How to cite this article: Orhan, H. & Yavşan, E. (2023). Artificial intelligence-assisted detection model for melanoma diagnosis using deep learning techniques. *Mathematical Modelling and Numerical Simulation with Applications*, 3(2), 159-169. <https://doi.org/10.53391/mmnsa.1311943>