Erzincan Üniversitesi
Fen Bilimleri Enstitüsü Dergisi
2024, 17(3), 870-883
ISSN: 1307-9085, e-ISSN: 2149-4584
Araştırma Makalesi

Erzincan University
Journal of Science and Technology
2024, 17(3), 870-883
DOI: 10.18185/erzifbed.1581918
Research Article

Enhanced Classification of Skin Lesions Using Fine-Tuned MobileNet and DenseNet121 Models with Ensemble Learning

Yasin SANCAR [1*]

[1]Ataturk University, Faculty of Open and Distance Education, Department of Information Systems, Erzurum, Türkiye

Received: 08/11/2024, **Revised:** 27/11/2024, **Accepted:** 30/11/2024, **Published:** 31/12/2024

Abstract

This study presents a deep learning approach for early detection of melanoma, one of the most dangerous skin cancers. In this article, all pre-trained models of the Keras library are trained with the ISIC skin cancer dataset available on Kaggle and the accuracy of each model is analyzed in detail. With the results obtained from the trained models, the models were fine-tuned to further optimize the performance of each model. After re-evaluation with fine-tuning, the accuracy rates were compared: DenseNet121 and MobileNet were found to be the two best models with high accuracy among the fine-tuned models. As such, these two models were combined in an ensemble approach to achieve a better overall accuracy. The skin cancer detection rate obtained with this ensemble approach is 93.03%. Therefore, the deep learning-based ensemble method appears to be a reliable and powerful technique that can be used to diagnose serious diseases such as skin cancer. This model can be used to provide a powerful support system with great potential to assist dermatologists in the early detection phase by easing workload and improving patient outcomes.

Keywords: Transfer learning, skin cancer classification, fine tuning, model ensemble

Topluluk Öğrenmesi ile İnce Ayarlı MobileNet ve DenseNet121 Modelleri Kullanılarak Cilt Lezyonlarının Geliştirilmiş Sınıflandırılması

Özet

Bu çalışma, en tehlikeli cilt kanserlerinden biri olan melanomun erken teşhisi için bir derin öğrenme yaklaşımı sunmaktadır Bu makalede, Keras kütüphanesinin önceden eğitilmiş tüm modelleri, Kaggle'da bulunan ISIC cilt kanseri veri kümesi ile eğitilmiş ve her modelin doğruluğu ayrıntılı olarak analiz edilmiştir Eğitilen modellerden elde edilen sonuçlarla, her modelin performansını daha da optimize etmek için modellere ince ayar yapılmıştır İnce ayar ile yeniden değerlendirme yapıldıktan sonra doğruluk oranları karşılaştırılmıştır DenseNet121 ve MobileNet, ince ayarlı modeller arasında yüksek doğruluk oranına sahip en iyi iki model olarak bulunmuştur Bu nedenle, bu iki model daha iyi bir genel doğruluk elde etmek için bir topluluk yaklaşımında birleştirilmiştir Bu topluluk yaklaşımı ile elde edilen cilt kanseri tespit oranı %93,03. Bu nedenle, derin öğrenme tabanlı topluluk yöntemi, cilt kanseri gibi ciddi hastalıkların teşhisinde kullanılabilecek güvenilir ve güçlü bir teknik olarak görünmektedir Bu model, iş yükünü hafifleterek ve hasta sonuçlarını iyileştirerek erken teşhis aşamasında dermatologlara yardımcı olmak için büyük potansiyele sahip güçlü bir destek sistemi sağlamak için kullanılabilir.

Anahtar Kelimeler: Transfer öğrenme, cilt kanseri sınıflandırma, ince ayar, model birleştirme

*Corresponding Author: yasinsancar@atauni.edu.tr
Yasin SANCAR, https://orcid.org/0000-0002-4200-1293

1. Introduction

Ultraviolet or ultraviolet rays with wavelengths between 100 and 400 nanometers have a significant effect on skin cancer [1]. These rays are divided into three categories: UV-A, UV-B and UV-C. UV-C rays are relatively less harmful than other rays and cannot reach the skin surface because they are absorbed by the atmosphere [2]. However, if UV-B and UV-C rays come into contact with the skin, it can cause damage to the skin tissue and cancer. Melanoma, a type of skin cancer, is divided into two types: benign and malignant, with an irregular structure containing several colors [3]. Early detection of malignant melanoma allows dermatologists to recommend surgical removal of the affected skin area to prevent the spread of the malformation in melanoma cells. Automatic disease recognition and diagnosis systems with machine learning and deep learning methods have been rapidly increasing in medical applications in recent years [4, 5, 6, 7]. These applications help specialists and significantly reduce their workload. Similarly, pre-diagnostic decision support systems have been proposed for skin cancer detection. In this study, a deep learning-based skin lesion detection mechanism is developed using images from the ISIC archive on Kaggle. In order to compare the performance of the proposed model with state-of-the-art models for early detection of malignant melanomas, only studies using the same dataset are evaluated.

1.1. Related studies

In a study by Basaran and C¸elik [8], the ISIC dataset was first trained with the EfficientNetB0 model and then deep features were obtained using Particle Swarm Optimization (PSO) and Genetic algorithm (GA) with the fully connected layer of this model. The features selected over different feature combinations were classified with Support Vector Machine, one of the classical machine learning methods, and an accuracy rate of 89.1% was achieved. Anand et al. [9] aimed to improve model accuracy by adding a flat layer, two dense layers with an activation function called LeakyRelu, and a sigmoid layer to a pre-trained VGG16 model and achieved 89.09% accuracy on the ISIC dataset. Sethanan et al. [10] classified melanoma, vascular lesions, melanocytic nevus, cutaneous fibromas, benign keratosis, and different carcinomas and skin moles using the HAM10000 dataset along with the ISIC dataset. In the proposed model, the input images are passed through the CNN model by applying image segmentation methods such as U-net, RP-Net, Threshold method, Edge detection and data augmentation such as rotation, shifting, and flipping in a dual artificial multiple intelligence system (AMIS). The proposed model outperformed the traditional CNN models with 98.4% on the hybrid dataset. Hussein et al. [11] applied various transfer learning networks such as AlexNet, ResNet-18, SqueezeNet, and ShuffleNet to the ISIC dataset and observed that the ResNet-18 model performed relatively better than other models with an accuracy of 89.9%. On the other hand, when precision, sensitivity and specificity values were compared, it was seen that the specificity rate exceeded 90% in the SqueezeNet model and surpassed the other models. Precision values showed slight differences in other models except ShuffleNet. However, when the F1 score value is analyzed, it is seen that the ResNet model has better performance than the other transfer learning models in general. Tuncer et al. [12] presented a CNN model called TurkerNet, which aims to improve classification performance by minimizing the number of trainable parameters by working on four basic components: input block, residual bottleneck block, efficient block and output. Since the proposed model shows high performance with an accuracy of 92.12% even with low trainable parameters, it can be frequently preferred in medical applications as a low-weight CNN model. Bazgir et al. [13] proposed an optimized Inception model for skin cancer based on the InceptionNet architecture with data augmentation and the addition of base layers. The proposed model was applied to the dataset from the ISIC archive and achieved 84.39% and 85.94% accuracy rates in Adam and Nadam optimizations, respectively. In another recent study presenting a deep learning-based approach, Prasad et al. [14] applied EfficientNet-B3, a deep transfer learning model, to ISIC data with rescaling, brightness, and contrast equalization preprocessing in the range of 20%. As a result of the experimental studies, 90.62% accuracy, 90.21% recall, 91.33% F1-score and 91.91% precision were obtained. A general comparison of the studies using the raw dataset I used in this study is presented in Table 1.

Table 1. Comparison of Classification Methods for Skin Lesion Analysis

Year	Study	Method	Classifier	Accuracy
2022	Basaran and Celik [8]	PSO-GA	SVM	89.1%
2022	Anand et al. [9]	VGG16	Softmax	89.09%
2022	Alfi et al. [15]	CNN	Softmax	92%
2023	Ramya and Sathiyabhama [16]	Enhanced genetic algorithm	SVM	89.19%
2023	Hussein et al. [11]	ResNet-18	Softmax	89.39%
2023	Shekar and Hailu [17]	DenseNet-169, local binary pattern	Random Forest	89.70%
2024	Bazgir et al. [13]	Optimized InceptionNet	Softmax	85.94%
2024	Prasad et al. [14]	EfficientNet-B3	Softmax	90.62%
2024	Turker et al. [12]	TurkerNet	Softmax	92.12%
	This study	**Ensemble of Fine Tuned MobileNet and DenseNet121**	**Softmax**	**93.03%**

1.2. Research Contributions

The main research contributions of this work are listed below: • The proposed MobileNet-DenseNet ensemble model outperforms other works in the literature using the same dataset, making significant progress in skin lesion diagnosis.

- The experimental results are presented in comparison with widely used pre-trained network models that have shown successful results in the literature.

- The highest performance was achieved by combining the DenseNet and MobileNet architectures from the pre-trained transfer learning models, and this achievement was compared with different studies in the literature using the same dataset.

- A fusion-based pre-trained transfer learning approach is proposed to improve skin lesion classification performance.

The flow diagram of the hybrid model combining DenseNet and MobileNet transfer learning networks with CNN is given in Figure 3.

2. Materials and Methods

2.1. Dataset

For the study, the Kaggle skin cancer dataset, an open-source platform containing 1800 images of benign and malignant skin lesions sized 224*224, was used [18]. This special dataset consists of training and test sets containing images divided into two classes: benign and malignant. Figure 1 shows a representative example of the dataset. Data augmentation included random rotations, zoom, and contrast adjustments to simulate diverse real-world scenarios, improving model robustness.

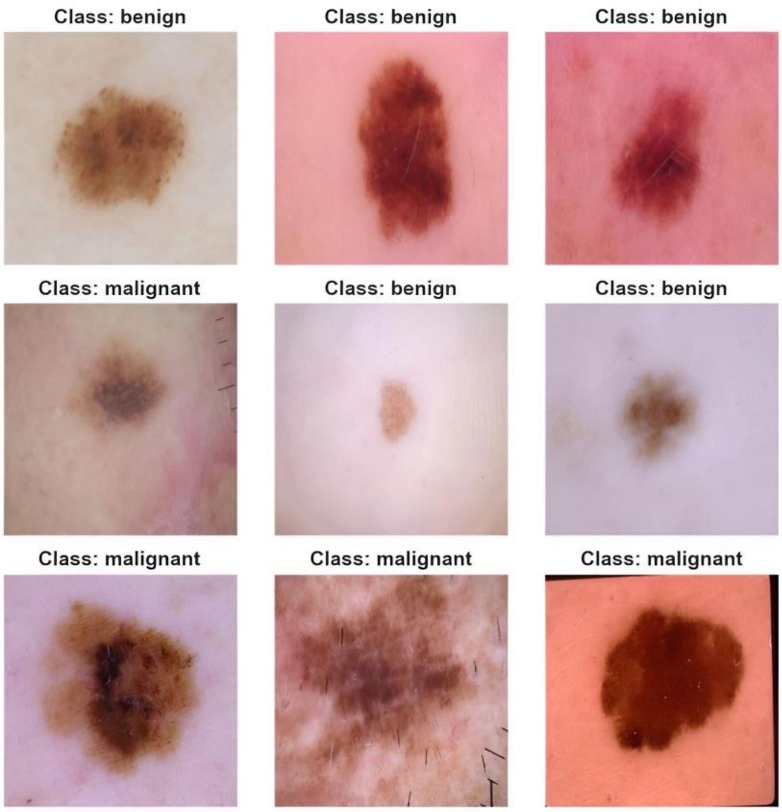

Figure 1. ISIC Archive sample skin lesions

As seen in Figure 2, the dataset of this study has been divided into training and test sets according to class labels. The benign dataset contains 360 images for testing and 1440 images for training. Similarly, for malignant training, 1197 training and 300 test images have been allocated.

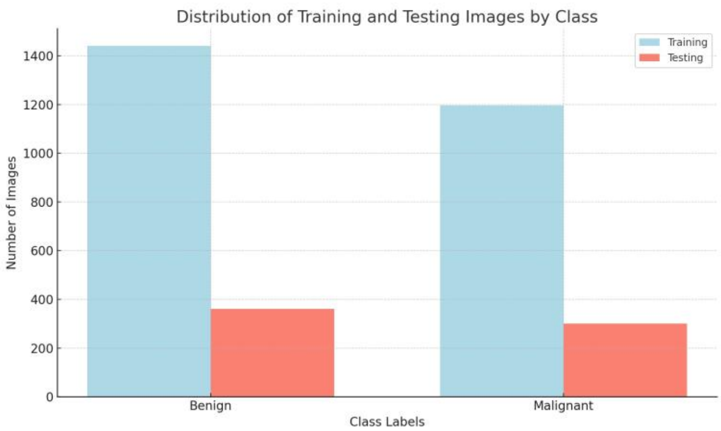

Figure 2. The number of benign and malignant lesions in the training and test dataset

2.2. The proposed ensemble transfer learning model

The use of pre-trained transfer learning models in the field of healthcare has significantly improved classification performance in recent years. The hyper-parameters of these models and the fact that the datasets on which the deep learning approaches are trained are balanced and contain a sufficient number of data affect the classification performance. In this study, I propose a hybrid use of a pre-trained deep learning approach on original data for skin lesion classification. Figure 3 shows the general flowchart of the proposed MobileNet-DenseNet ensemble model. In the proposed model, all layers in MobileNet are made trainable and fine tuned without any preprocessing of the original data set. Then, using the CNN architecture, GlobalAverageMaxPooling2d, Dense 1024, Batch Normalization, Dropout (0.5), Dense 64 and Dense 2 layers were used to reduce the number of classes to two, malignant and benign, by gradually reducing the layers in order not to lose the features obtained from the pre-trained model. As a second model, the last 171 layers of the DenseNet121 model out of a total of 427 layers are trainable and subjected to fine tuning. After the same CNN operations, the output obtained from both models was combined end-to-end to obtain a hybrid ensemble model.

The Global Average Pooling layer is a process that calculates the average output of the feature map in the previous layer. As shown in Figure 3, a

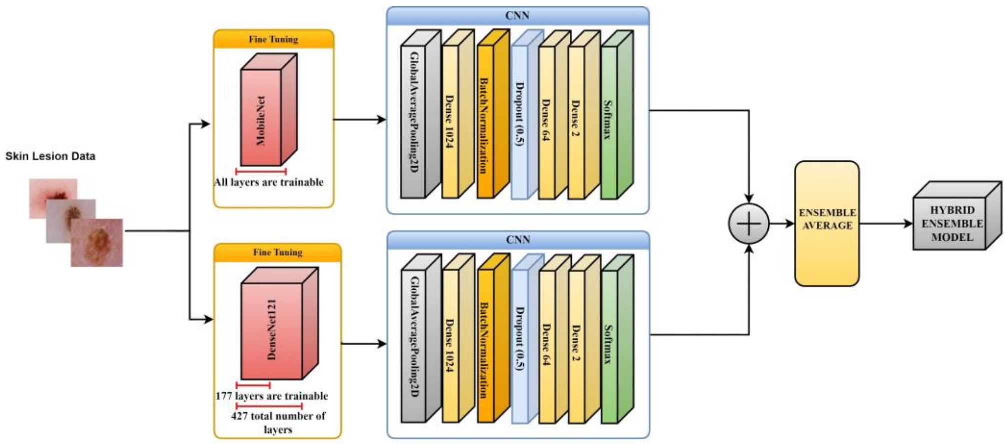

Figure 3. The diagram of the proposed ensemble transfer learning model

Global Average Pooling layer preceding the Dense layer fully connected layer is used to extract features from the trainable layers where fine tuning is performed. With this average calculation, the features are significantly reduced, preparing the model for the final classification layer. In Global Average Pooling, overfitting is avoided by averaging the feature map. A dropout layer is used to reduce overfitting during the training process. It is the elimination of some memorizing nodes in the network to prevent the network from being memorized. Thus, the memorization of the network is tried to be eliminated. The dropout layer is a flattening layer for fully connected layers. Dropout increases the smoothing ability of the neural network. With dropout, neurons in the network are randomly assigned a zero weight value. For this process, the dropout rate is set to 0.5 to make the model robust to small changes in the input and to ensure high performance models. Dense fully connected layers, which are added after the dropout process, are gradually reduced to 2 classes in order to prevent the loss of features and these features are given to the Softmax layer to give the classification result. The training process employed the Adam optimizer with a learning rate of 0.001, chosen for its balance between convergence speed and stability. Other hyperparameters, such as batch size (32) and dropout rates (0.5), were fine-tuned through cross-validation to avoid overfitting.

2.2. Mathematical Model for Global Average Pooling, Dropout, and Average Ensemble

Let F represent the output feature map of the preceding layer, as follows:

$$F = \begin{bmatrix} f_{1,1} & f_{1,2} & \cdots & f_{1,n} \\ f_{2,1} & f_{2,2} & \cdots & f_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{m,1} & f_{m,2} & \cdots & f_{m,n} \end{bmatrix}$$

where the output of the *j*-th channel of the *i*-th feature map is represented by $f_{i,j}$.

The average output for every channel is determined by the Global Average Pooling operation:

$$G = \frac{1}{H \times W} \sum_{h=1}^{H} \sum_{w=1}^{W} F_{h,w}$$

where H and W represent the feature map’s height and width, respectively, to create a vector G that represents the feature map’s summary. This technique works well to lessen overfitting. [19]

In order to avoid overfitting, the Dropout layer then randomly changes a fraction p of the input units to zero during training:

$$Y = D(F) = \{ \begin{matrix} F & \text{with probability } 1 - p \\ 0 & \text{with probability } p \end{matrix}$$

This technique helps to improve the generalization of neural networks [20].

Finally, the Dense layer receives the output from the Dropout layer and uses C units for classification:

$$Z = W \cdot Y + b$$

where, prior to applying the Softmax activation function, Z is the output, b is the bias vector, W is the weight matrix, and Y is the input from the Dropout layer:

$$P = \text{Softmax}(Z)$$

where P gives the predicted probabilities for each class.

To employ an average ensemble strategy to integrate many models, let M represent the number of models and Oi represent the output of the i-th model:

$$O_i = \text{Softmax}(Z_i)$$

The final ensemble output Oensemble can be computed as the average of the outputs from all models:

$$O_{\text{ensemble}} = \frac{1}{M}\sum_{i=1}^{M} O_i$$

This averaging method is a common ensemble strategy used to enhance predictive performance. [21].

3. Experimental Results

In the study, the dataset was initially trained on pre-trained models and accuracy rates were obtained. These models were first tested without any modifications and the accuracy rates for each were determined. After the initial training phase, fine-tuning was applied to each model to further optimize model performances. Fine-tuning was used to adjust model weights in pre-trained models to improve their ability to generalize from training data. After fine-tuning, the models were trained and the accuracy rates were redetermined. After this step, the DenseNet121 and MobileNet models had the highest accuracy among the fine-tuned models. Table 2 shows the 5 models with the highest accuracy. Fine-tuning adjusts pre-trained model parameters to better adapt to specific datasets, improving performance on new tasks.

Table 2. Individual performance of transfer learning models after fine tuning and CNN applications

Model	Precision	Recall	F1 Score	Accuracy	Epoch
MobileNet	0.9242	0.9242	0.9241	0.9242	22
DenseNet121	0.9232	0.9227	0.9228	0.9227	68
ResNet50V2	0.8803	0.8803	0.8800	0.8803	63
VGG19	0.8787	0.8787	0.8787	0.8787	49
Xception	0.8713	0.8712	0.8709	0.8712	68

The graphs of training and test accuracy, training and validation loss values of the DenseNet121 model according to the number of epochs are given in Figure 4. As can be seen from the figure, while the training accuracy increases rapidly when the number of epochs is increased up to 20, it reaches a more regular accuracy rate after 60 epochs. Similarly, the test accuracy varied at each step, but it reached over 90% after the 60th epoch. When comparing the training and validation loss values, the training loss value decreased in the opposite direction of the accuracy rate as the number of epochs increased and progressed more consistently than the test loss.

Figure 4. Training and Test Accuracy/Loss graphs for DenseNet121 according to the number of epochs

Figure 5 shows the training-test accuracy and training-validation loss graphs of the MobileNet model according to the number of epochs. Similar to the DenseNet graph, while more stable accuracy and loss rates are determined on the train data, large differences are observed between 20 and 40 epochs in the test set. Therefore, the number of epochs was truncated at 22 to minimize test loss.

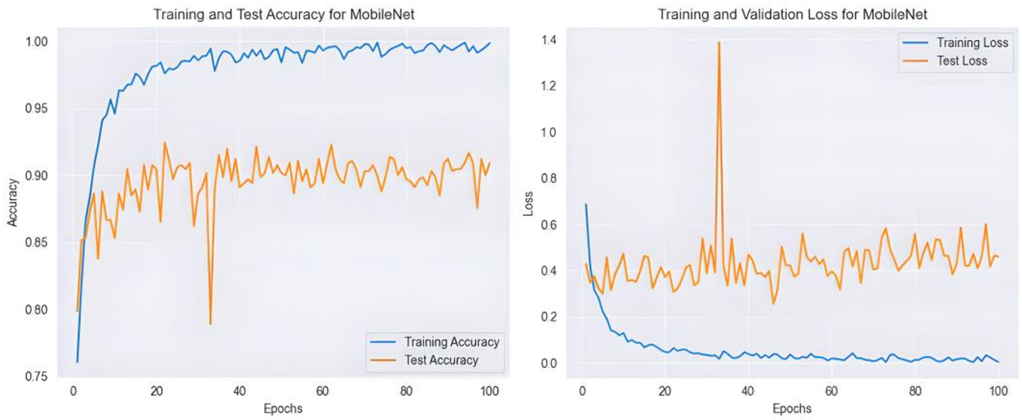

Figure 5. Training and Test Accuracy/Loss graphs for MobileNet according to the number of epochs

An ensemble approach was used to further improve the accuracy of the classification system. Ensemble learning involves combining the predictions of multiple models to improve overall performance and robustness. In the study, the fine-tuned outputs of several transfer learning models were combined together to create a more comprehensive and accurate classification mechanism. The ensemble method exploits the unique strengths of each model, reducing the weaknesses that any one model may have. This method reduced the probability of misclassification and increased the model’s ability to generalize across different data samples. The result is a significant improvement in accuracy and reliability, as demonstrated by the MobileNet and DenseNet121 ensemble, which reached an accuracy of 93.03%. Table 3 shows the results obtained by combining the models with the highest accuracy with the ensemble method. For model ensembling, the average method was used, where the outputs of the individual models, MobileNet and DenseNet121, were averaged to make the final prediction. This approach leverages the complementary strengths of the models: MobileNet's efficiency and DenseNet121's ability to extract deep features. By averaging their outputs, the ensemble model reduces the impact of potential biases or weaknesses of individual models, leading to improved robustness and accuracy. This method ensures that each model contributes equally to the final decision, making it an effective and computationally efficient strategy for combining predictions in a classification task.

Table 3. The performances of ensemble models

Model	*Precision*	*Recall*	*F1 Score*	*Accuracy*
MobileNet + DenseNet121	0,9307	0,9303	0,9303	0,9303
MobileNet + Xception	0,9218	0,9212	0,9209	0,9212
DenseNet121 + VGG19	0,9148	0,9136	0,9137	0,9136
MobileNet + ResNet50V2	0,9121	0,9121	0,9121	0,9121
MobileNet + VGG19	0,9120	0,9121	0,9120	0,9121
DenseNet121 + ResNet50V2	0,9120	0,9121	0,9120	0,9121
DenseNet121 + Xception	0,9030	0,9030	0,9030	0,9030
ResNet50V2 + VGG19	0,8984	0,8984	0,8984	0,8984
Xception + VGG19	0,8940	0,8939	0,8937	0,8939
ResNet50V2 + Xception	0,8842	0,8833	0,8828	0,8833

In Figure 6, the performance of DenseNet121, MobileNet and DenseNet121+ MobileNet hybrid transfer learning model on skin cancer is measured classwise in confusion matrices. In distinguishing two very similar classes, the total number of misclassifications of the DenseNet121 model is 51 while the number of misclassifications of MobileNet is 50. The total number of misclassifications obtained as a result of the combination of both models is reduced to 46. It can be said that the proposed ensemble model is effective in classifying data that are very similar to each other.

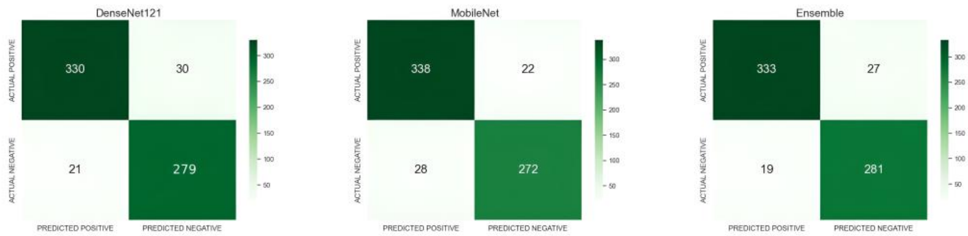

Figure 6. MobileNet, DenseNet121 and proposed model's confusion matrices

4. Discussion

The results of this study show that the performance of pre-trained deep learning models in the classification of skin lesions is enhanced by the ensemble technique. By fine-tuning the MobileNet and DenseNet121 models and then creating an ensemble, superior performance metrics have been achieved compared to individual models or other ensemble combinations. The ensemble model combining MobileNet and DenseNet121 has achieved the highest precision, recall, F1 score, and accuracy, surpassing other commonly used models such as ResNet, VGG19, and Xception. One of the most significant strengths of this approach is its hybrid structure that combines the strengths of both MobileNet and DenseNet. MobileNet is known for its efficiency and speed, complementing DenseNet's ability to leverage deeper feature representations. While fine-tuning the last 171 layers of DenseNet121, allowing all layers in MobileNet to be trainable ensured that both models adapted to the unique characteristics of the skin lesion dataset without overfitting. The use of dropout layers and global average pooling also contributed to the overall performance of the model. These techniques reduced overfitting and ensured that the model memorized the training data, which is crucial in medical imaging tasks where generalization to new, unseen data is vital. Additionally, instead of relying on a simple voting scheme, the end-to-end combination of the models' outputs ensured the full utilization of each model's strengths and provided a robust classification output. Experimental results further emphasize that independent models like ResNet50V2 and Xception perform reasonably well, but they cannot distinguish between benign and malignant skin lesions as effectively as the ensemble model. Confusion matrices reveal that the combination of MobileNet and DenseNet121 reduces misclassification errors, especially when distinguishing between classes with very similar features, compared to individual models. This is very important in the field of healthcare, where accurate classification can lead to significant clinical outcomes.

Despite these promising results, there are some limitations in our approach. For example, while our ensemble model achieved high performance metrics, the computational complexity and training time were higher compared to individual models. Especially the DenseNet121 model required more epochs to stabilize compared to MobileNet, indicating that future studies should explore more efficient ways to combine deep learning models so that the computational load does not increase significantly.

5. Conclusion

This model, which proposes a hybrid ensemble model combining MobileNet with DenseNet121, was used for the classification of skin lesions. The approach here was to leverage the strengths of both models. By fine-tuning the final layers of the models, better learning was achieved. Subsequently, high accuracy was achieved using the merging method. Thanks to this, it has reached the latest technological performance in classifying skin lesions as benign or malignant. It has been determined that such hybrid approaches can provide better results compared to individual models obtained from transfer learning and other ensemble combinations. This study demonstrates the potential of transfer learning in healthcare applications, particularly in the classification of skin lesions. The reduction in misclassifications with the collective approach underscores its value, especially in a clinical setting where accurate diagnosis is paramount. However, further research is needed on the optimal computational efficiency of this approach, especially for its use in real-time applications. Large-scale and more diverse datasets will further validate the generalization of the proposed model. These promising results enable the investigation of community methods on medical images in many other applications where appropriate and efficient classification, which is absolutely necessary for the clinical decision-making process, is required. The proposed ensemble model significantly reduced misclassification rates compared to individual models, demonstrating its potential in medical imaging. By leveraging MobileNet's efficiency and DenseNet121's deeper feature extraction, the ensemble achieved higher accuracy and generalization.

Ethics in Publishing

There are no ethical issues regarding the publication of this study

Author Contributions

Yasin SANCAR: Designed and wrote the manuscript, carried out the experimental studies, conclusions and interpretation of the experiment.

References

[1] Nisar, D.M., Amin, R., Shah, N.H., Al Ghamdi, M. A., Almotiri, S. H., & Alruily, M. (2021). Healthcare techniques through deep learning: issues, challenges and opportunities. IEEE Access, 9, 98523–98541.

[2] Alfi, I. A., Rahman, M. M., Shorfuzzaman, M., & Nazir, A. (2022) A non-invasive interpretable diagnosis of melanoma skin cancer using deep learning and ensemble stacking of machine learning models. Diagnostics, 12(3), 726.

[3] Anand, V., Gupta, S., Koundal, D., & Singh, K. (2023) Fusion of U-Net and CNN model for segmentation and classification of skin lesion from dermoscopy images. Expert Systems with Applications, 213, 119230.

[4] Başaran, E., & Çelik, Y. (2022) Skin cancer diagnosis using CNN features with Genetic Algorithm and Particle Swarm Optimization methods. Transactions of the Institute of Measurement and Control, 01423312241253926.

[5] Bazgir, E., Haque, E., Maniruzzaman, M., & Hoque, R. (2024) Skin cancer classification using Inception Network. World Journal of Advanced Research and Reviews, 21(2), 839–849.

[6] Dimen, L., Tahâs, V.-S., Borşan, T., & Ferencz, Z. (2016) Ultraviolet Rays and Its Environmental Effects. International Multidisciplinary Scientific GeoConference: SGEM, 1, 787–792.

[7] Hussein, Z. J., Hussein, A. M., Maki, G. I., & Gheni, H. Q. (2023) Improved model for skin illnesses classification utilizing gray-level co-occurrence matrix and convolution neural network. Journal of Advances in Information Technology, 14(6), 1273–1279.

[8] Ker, J., Wang, L., Rao, J., & Lim, T. (2017) Deep learning applications in medical image analysis. Ieee Access, 6, 9375–9389.

[9] Kilic, U., Karabey Aksakalli, I., Tumuklu Ozyer, G., Aksakalli, T., Ozyer, B., & Adanur, S. (2023) Exploring the Effect of Image Enhancement Techniques with Deep Neural Networks on Direct Urinary System (DUSX) Images for Automated Kidney Stone Detection. International Journal of Intelligent Systems, 2023(1), 3801485.

[10] Latif, J., Xiao, C., Tu, S., Rehman, S. U., Imran, A., & Bilal, A. (2020) Implementation and use of disease diagnosis systems for electronic medical records based on machine learning: A complete review. IEEE Access, 8, 150489–150513.

[11] Mohania, D., Chandel, S., Kumar, P., Verma, V., Digvijay, K., Tripathi, D., Shah, D. (2017) Ultraviolet radiations: Skin defense-damage mechanism. Ultraviolet Light in Human Health. Diseases and Environment, 71–87.

[12] Sethanan, K., Pitakaso, R., Srichok, T., Khonjun, S., Thannipat, P., Wanram, S., … Others. (2023) Double AMIS-ensemble deep learning for skin cancer classification. Expert Systems with Applications, 234, 121047.

[13] Shekar, B. H., & Hailu, H. (2023) Fusion of features extracted from transfer learning and handcrafted methods to enhance skin cancer classification performance. In Computer Vision and Machine Intelligence: Proceedings of CVMI 2022, Springer, 243–257.

[14] Tumpa, P. P., & Kabir, M. A. (2021) An artificial neural network based detection and classification of melanoma skin cancer using hybrid texture features. Sensors International, 2, 100128.

[15] Tuncer, T., Barua, P. D., Tuncer, I., Dogan, S., & Acharya, U. R. (2024) A lightweight deep convolutional neural network model for skin cancer image classification. Applied Soft Computing, 111794.

[16] Panneerselvam, R., Balasubramaniam, S. (2023) Multi-Class Skin Cancer Classification using a hybrid dynamic salp swarm algorithm and weighted extreme learning machines with transfer learning. Acta Informatica Pragensia, 12(1), 141–159.

[17] Prasad, C. R., Bilveni, G., Priyanka, B., Susmitha, C., Abhinay, D., & Kollem, S. (2024) Skin Cancer Prediction using Modified EfficientNet-B3 with Deep Transfer Learning. 2024 IEEE International Conference for Women in Innovation, Technology & Entrepreneurship (ICWITE), 519–523.

[18] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014) Dropout: a simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research, 15(1), 1929–1958.

[19] Dietterich, T. G. (2000) Ensemble methods in machine learning. International Workshop on Multiple Classifier Systems, Springer, 1–15.

[20] Lin, M. (2013) Network in network. arXiv Preprint arXiv:1312. 4400.

[21] Fanconi, C. (2019) Skin cancer: malignant vs. benign-processed skin cancer pictures of the ISIC archive.