Korkut, Ş. G., Kocabaş, H., Kurban, R. (2024). A Comparative Analysis of Convolutional Neural Network Architectures for Binary Image Classification: A Case Study in Skin Cancer Detection. *The Black Sea Journal of Sciences*, 14(4), 2008-2022.

The Black Sea Journal of Sciences, 14(4), 2008-2022, 2024. DOI: 10.31466/kfbd.1515451

KFBD

Karadeniz Fen Bilimleri Dergisi
The Black Sea Journal of Sciences
ISSN (Online): 2564-7377 https://dergipark.org.tr/tr/pub/kfbd

Araştırma Makalesi / Research Article

A Comparative Analysis of Convolutional Neural Network Architectures for Binary Image Classification: A Case Study in Skin Cancer Detection

Şerife Gül KORKUT[1], Hatice KOCABAŞ[2], Rifat KURBAN[3*]

Abstract

In this study, a comprehensive comparative analysis of Convolutional Neural Network (CNN) architectures for binary image classification is presented with a particular focus on the benefits of transfer learning. The performance and accuracy of prominent CNN models, including MobileNetV3, VGG19, ResNet50, and EfficientNetB0, in classifying skin cancer from binary images are evaluated. Using a pre-trained approach, the impact of transfer learning on the effectiveness of these architectures and identify their strengths and weaknesses within the context of binary image classification are investigated. This paper aims to provide valuable insights for selecting the optimal CNN architecture and leveraging transfer learning to achieve superior performance in binary image classification applications, particularly those related to medical image analysis.

Keywords: Convolutional Neural Networks (CNNs), Transfer Learning, Binary Image Classification, CNN Architecture Comparison, Skin Cancer Detection.

İkili Görüntü Sınıflandırma için Evrişimsel Sinir Ağı Mimarilerinin Karşılaştırmalı Analizi: Cilt Kanseri Tespitinde Bir Vaka Çalışması

Öz

Bu çalışmada, ikili görüntü sınıflandırması için Evrişimsel Sinir Ağı (CNN) mimarilerinin kapsamlı bir karşılaştırmalı analizi sunulmuş ve transfer öğreniminin faydalarına vurgu yapılmıştır. MobileNetV3, VGG19, ResNet50 ve EfficientNetB0 gibi önde gelen CNN modellerinin ikili görüntülerden cilt kanseri sınıflandırmadaki performans ve doğruluğu değerlendirilmiştir. Önceden eğitilmiş bir yaklaşım kullanılarak, transfer öğreniminin bu mimarilerin etkinliği üzerindeki etkisi araştırılmış ve ikili görüntü sınıflandırması bağlamında güçlü ve zayıf yönleri belirlenmiştir. Bu makale, optimal CNN mimarisinin seçimi ve transfer öğreniminden yararlanarak ikili görüntü sınıflandırma uygulamalarında, özellikle tıbbi görüntü analiziyle ilgili olanlarda, üstün performans elde etme konusunda değerli içgörüler sağlamayı amaçlamaktadır.

Anahtar Kelimeler: Evrişimsel Sinir Ağları (CNN'ler), Transfer Öğrenimi, İkili Görüntü Sınıflandırma, CNN Mimari Karşılaştırması, Cilt Kanseri Tespiti.

[1,2,3]Abdullah Gül University, Department of Computer Engineering, 38080, Kayseri, Türkiye,
serifegul.korkut@agu.edu.tr hatice.kocabas@agu.edu.tr rifat.kurban@agu.edu.tr

[*]Sorumlu Yazar/Corresponding Author
Geliş/Received: 12.07.2024 **Kabul/Accepted:** 14.10.2024 **Yayın/Published:** 15.12.2024

1. Introduction

Skin cancer, comprising melanoma and non-melanoma skin cancers, is a pervasive and increasingly prevalent form of cancer, primarily arising from melanocytes and epidermal cells (Craythorne and Al-Niami, 2017). The urgent need for accurate and efficient skin cancer detection methods drives this study, which explores the transformative potential of transfer learning for binary image classification. Given the high incidence of skin cancer and its potential impact on public health, the utilization of CNN architectures with transfer learning has shown promise in improving diagnostic accuracy, as demonstrated in numerous studies, including the work by Dildar et al. (2021).

Conventional skin cancer detection techniques typically depend on dermatologists manually examining patients; in this case, the keen eyes of experienced professionals are vital in spotting possible lesions. However, the development of artificial intelligence, particularly CNNs, is causing a paradigm shift in dermatological diagnosis, as it does in many other fields. This technological advancement creates opportunities for automated analysis of dermatoscopic images, showing promise for more accurate and reliable skin lesion detection.

The challenge in skin cancer detection extends beyond simply designing robust CNN architectures. The complexity lies in effectively refining their performance for accurate binary image classification. While CNNs offer a systematic and objective approach to analyzing dermatoscopic images, surpassing the subjectivity of traditional methods, optimizing their capabilities for accurate cancer identification remains a key focus in this evolving field. However, these CNNs' efficacy isn't only a function of how sophisticated their architecture is; it also has to do with how well they adjust to the unique characteristics of dermatological photos. The focus of binary classification for skin cancer diagnosis is on differentiating critical characteristics that clearly indicate the presence or absence of malignancy.

The primary objective of this paper is to assess different CNN architectures' performance while taking transfer learning's subtle effects into account. We want to know how well models like VGG-19, ResNet50, MobileNetV3, AlexNet, and EfficientNetB0 adapt to the complexity of skin cancer binary classification by using pre-trained weights. The objective of this research is to expose the innate advantages and disadvantages of every design, offering valuable perspectives on their suitability and efficiency in practical situations.

The significance of this study lies in its potential to identify optimal CNN architectures for binary image classification in skin cancer diagnosis. It is also anticipated that this research will elucidate the strategic application of transfer learning to enhance the performance of these systems, especially in scenarios characterized by limited datasets. By conducting a comparative analysis of models trained on the "Skin Cancer Binary Classification Dataset," valuable insights are sought that are expected to shape the future trajectory of automated skin cancer detection.

The rest of the paper is organized as follows: section 2 gives a brief summary of the literature, section 3 describes the dataset, and the CNN methods used, section 4 gives the results and the final concluding marks are given in section 5.

2. Related Works

CNNs have emerged as a powerful tool for image classification, demonstrating exceptional ability to automatically extract relevant features from input data (Ullah & Mahmoud, 2021). Within the domain of medical image analysis, CNNs have shown significant promise for automating diagnostic tasks, including skin cancer detection (Dildar et al., 2021).

Previous research has explored the performance of various CNN architectures for binary image classification, including prominent models such as ResNet, MobileNet and EfficientNet (Tan, 2019). These studies highlight the importance of selecting architectures that balance accuracy with computational efficiency, particularly when applied to resource-constrained settings (Sobczak & Kapela, 2022).

Transfer learning has further enhanced the capabilities of CNNs by allowing models to leverage knowledge gained from previous training on large datasets (Prima & Bouhorma, 2020). This approach has proven particularly valuable in medical image analysis, where datasets are often limited, and has been successfully applied to classify Alzheimer's disease stages using neuroimaging data (Tufail et al., 2021). The use of pre-trained models can significantly improve the performance of CNNs in tasks like anomaly detection and classification, which has implications for fields like malware detection and cognitive radio (Suciu et al., 2019; Geng et al., 2022).

Despite these advancements, challenges remain in optimizing CNN architectures for specific medical image analysis tasks, particularly for accurate and robust skin cancer detection. Existing research often focuses on general image classification tasks, such as hyperspectral image classification (Bai et al., 2019), or utilizes datasets that may not fully represent the complexity and variability of dermatoscopic images.

Transfer learning techniques are widely adopted due to the limited size of medical datasets. Abdulridha and Savaş (2022) employed data augmentation to mitigate the class imbalance in the ISIC dataset, significantly enhancing model performance. Their DenseNet121 model reached a 99.6% accuracy, surpassing other state-of-the-art techniques like ResNet-50 and EfficientNet, which achieved accuracies of 98.1% and 98.5%, respectively. Similarly, Islam and Panta (2024) applied five transfer learning approaches, including ResNet-50, MobileNet, InceptionV3, DenseNet-169, and

InceptionResNetV2, to the ISIC dataset, with ResNet-50 yielding an accuracy of 93.5%.They utilized data augmentation and fine-tuned transfer learning models, where ResNet-50 again stood out with a precision of 0.94 and an F1 score of 0.86. Rashid et al. (2022) applied MobileNetV2 to classify melanoma and benign lesions using the ISIC-2020 dataset, achieving an accuracy of 98.2%.

This research addresses these limitations by conducting a comprehensive comparative analysis of prominent CNN architectures specifically tailored for binary image classification in skin cancer detection. We utilize a specialized dataset and investigate the impact of transfer learning on model performance, seeking to identify optimal architectures and strategies for achieving superior accuracy in this critical diagnostic domain.

3. Material and Methods

The method used to classify skin cancer using the Kaggle Skin Cancer Binary Classification Dataset is described in this section. A total of five different CNN architectures—VGG19, MobileNetV3, AlexNet, ResNet50, and EfficientNetB0—were included in the comparison. These previously trained models were adjusted and assessed on different test sets using transfer learning. Furthermore, effective model training was enabled by Kaggle's cloud-based architecture with T4X2 GPU acceleration. By employing these approaches, the research seeks to improve the models' capacity to identify patterns in skin cancer, with an emphasis on effectiveness, efficiency, and generalization over a range of complexity.

3.1. Dataset Description

The Kaggle Skin Cancer Binary Classification Dataset was used in this study. The dataset consists of a total of 288 skin lesion images, evenly distributed between the two classes. There are 144 images labeled as "Cancer" and 144 images labeled as "Non-Cancer."

There are two primary subdirectories within the dataset organization:

1. Cancer Class Images:
    - Training Images: cancer/training/
    - Testing Images: cancer/testing/
    - Total Images: 144
2. Non-Cancer Class Images:
    - Training Images: non_cancer/training/
    - Testing Images: non_cancer/testing/
    - Total Images: 144

The JPEG format ensures a consistent representation for model compatibility across all pictures. The image distribution between the 'Cancer' and 'Non-Cancer' classes is balanced with an equal number of images in each, which is essential to avoid biases during model training and evaluation.

The dataset was preprocessed using many procedures in order to make it ready for model training: For every class, the original data set was divided into testing and training sets. In order to assess the model's generalization ability, this part enables it to be trained on an independent set of pictures and assessed on an independent set. During pre-processing, each class's picture file names were randomly shuffled. By doing this, inadvertent biases that may result from categorizing the photos are reduced. Keras ImageDataGenerator was used to apply data augmentation to the training set, producing augmented images through arbitrary transformations like rotation, flipping, and zooming. This process exposes the model to a wider range of variables, enhancing its generalization capacity.

3.2. CNN Architectures for Comparison

Figure 1 illustrates the process of transfer learning using a pre-trained CNN, which enables a model to adapt to a new task by leveraging insights from prior training. It delineates the process into distinct phases: first, input images for model training are introduced, either pre-processed or raw; second, a pre-trained CNN, potentially from large datasets like ImageNet, transfers its parameters to the initial layers of the new model. Subsequently, fine-tuning adjusts these parameters for the specific task through iterative training with the new dataset. The model's efficacy is then evaluated using metrics like accuracy and precision, leading to the reporting of performance and findings. Additionally, the diagram underscores the versatility of employing diverse pre-trained CNN models such as MobileNetV3, VGG19, EfficientNetB0, ResNet50, and AlexNet, across various tasks, including cancer classification.

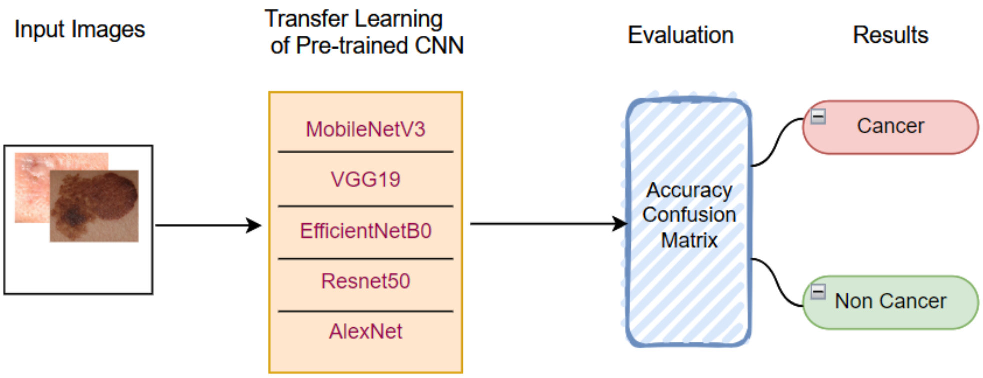

Figure 1. General architecture of the proposed research study.

3.2.1 VGG19

VGG19, a 19-layer Deep Convolutional Neural Network architecture, is known for its structural simplicity and efficient training on large datasets due to its numerous parameters (Li et al., 2022). It primarily comprises completely linked layers after a series of sequentially stacked, small-sized (3x3) convolutional layers.

Its construction is simple structurally. Its tiny depth and core size can improve learning capacity. Because there are a lot of parameters, it can be trained efficiently on big data sets.

Numerous applications, including the automated diagnosis of retinopathy of prematurity, have made extensive use of VGG19 (Huang et al., 2020). Furthermore, for better performance, VGG19 has been used in conjunction with other models. As an example, ensemble networks that include several VGG19 designs have demonstrated great accuracy in the identification of diabetic retinopathy (Hasan & Aleef, 2019). Also, VGG19 has been effectively incorporated into a number of architectures to improve voice processing. As an illustration of the potency of merging these models, (Kashani et al., 2019) integrated VGG19 into the VGG19-UNet architecture for voice augmentation. Furthermore, Jia & Li (2022) demonstrated the adaptability of VGG19 in various applications by using it in AE-VGG19 models for feature extraction. These integrations demonstrate how versatile and effective VGG19 is at improving speech-related tasks by combining creative architectural designs.

3.2.2 MobileNetV3

MobileNetV3, a specialized architecture renowned for its efficient and light design, is geared for mobile and edge device performance (Sandler et al., 2018). It stands out for being straightforward, having little processing overhead, and utilizing cutting-edge methods like inverse residuals and linear bottlenecks. The model's activation and input preprocessing properties amplify its distinct qualities. MobileNetV3 is well-suited for on-device inference on devices with constrained resources since it integrates modules such as Inverted Bottleneck blocks to develop inventive designs like Self-Attention MobileNet (Garg et al., 2021).

It is a simple model with few parameters and a cheap computing cost. It uses cutting-edge methods including linear bottlenecks and inverse residuals. Its input preprocessing and activation features make it noteworthy.

In addition, MobileNetV3 integrates the h_swish activation function and the SE attention mechanism to improve network performance, especially in situations where computing resources are few (Guo-zhan et al., 2023). Tasks including vehicle-pedestrian detection and skin disease categorization have demonstrated its efficiency and flexibility, demonstrating its versatility across various applications (Hu et al., 2022; Deng & Wu, 2022).

3.2.3 AlexNet

After winning the 2012 ImageNet competition, AlexNet, a groundbreaking deep learning architecture, has had a major impact on computer vision research (Russakovsky et al., 2015). Compared to earlier models, this architecture's eight layers provided a deeper structure and larger convolutional filters, which made it possible to learn more complicated features. Its capabilities were further improved by the use of Rectified Linear Unit (ReLU) activation functions and Local Response Normalization (LRN) layers (Tinnathi & Sudhavani, 2022).

Compared to earlier models, a deeper structure and larger size convolutional filters enable the learning of more complex features. It makes use of Local Response Normalization (LRN) layers and ReLU activation functions.

Additionally, AlexNet has been improved and changed in many research to increase its effectiveness in particular activities. To improve feature extraction capabilities, for example, researchers have suggested adding new layers and modules (Xu et al., 2021); talking about how different preprocessing techniques and convolution kernel sizes affect model performance (Bu et al., 2022); and using AlexNet as a feature extractor in conjunction with other algorithms, such as SVM, for classification tasks (Al-Mekhlafi et al., 2022).

3.2.4 ResNet50

ResNet50 is a 50-layer deep neural network design that is well-known for using residual blocks to make deep neural network training easier. Identity Mapping requires features like small-sized convolutions and transition layers, which are included in the architecture of ResNet50 (Mvoulana et al., 2021). Because they solve the problem of disappearing gradients, these residual blocks are essential in facilitating the development of deeper networks (Liu et al., 2022).

Deeper networks may be created thanks to the connections found in residual blocks, which lessen the issue of vanishing gradients. Features like small-sized convolutions and transition layers are part of Identity Mapping.

In order to improve feature extraction and resilience, researchers have also improved the ResNet50 model by adding more modules like the convolutional block attention module (CBAM) (Du et al., 2023). Furthermore, research has demonstrated that ResNet50's convolutional layers' depth is adequate for completely extracting visual features, improving performance on tasks like image classification (Wang et al., 2022).

3.2.5 EfficientNetB0

The base model of the EfficientNet series, EfficientNetB0, is renowned for its balanced construction and composite scaling approach that maximizes depth, width, and resolution (Kamiri et al., 2022). This model's efficiency in terms of both classification accuracy and computing complexity has led to its widespread implementation in a variety of applications (Laschowski et al., 2021).

The model's structure is balanced, and its depth, breadth, and resolution are all increased. A good place to start for advanced transfer and feature learning.

Its application in the diagnosis of retinal diseases and the prediction of picture credibility in the identification of fake news highlights its wide range of applications and influence in several fields (Singh & Sharma, 2021).

3.3 Transfer Learning Implementation

Using the procedures listed below, five CNN transfer learning models were developed and applied to the dataset.

1. A Dense layer with two output units that corresponded to the binary classification job was used to replace the final classification layer of each pre-trained model.

2. During training, all of the base models' layers aside from the recently added Dense layers were frozen to protect previously learned data.
3. Models were trained with categorical cross-entropy loss and assembled using the Adam optimizer with a learning rate of 0.00001. Models were assessed on a different test set after training on enhanced training data.

Through the use of these transfer learning methodologies, models are better able to distinguish between patterns associated with skin cancer within the particular dataset utilized in this work by using characteristics from a varied dataset like ImageNet.

In conclusion, this approach used a variety of CNN architectures with transfer learning to classify skin cancer cases. The method took performance and efficiency into account at different model complexity levels. To give a more reliable assessment, every model was also trained and evaluated thirty times, with the results averaged. The models' resilience and capacity for generalization were intended to be enhanced by the chosen preprocessing methods and transfer learning approaches.

3.4 Experimental Setup

Using the T4X2 GPU that Kaggle offered, we were able to utilize its cloud-based architecture to harness the computational capacity needed to train deep neural networks, greatly accelerating the model training process. Additionally, it made optimal use of processing resources and allowed for quicker iterations. Kaggle Notebooks, a platform that gives users access to powerful GPUs, was used to conduct the research, with the T4X2 GPU providing substantial processing power for the quick training of intricate deep learning models.

4. Result and Discussion

This section presents a thorough comparison study of CNN architectures for binary image classification tasks, with a focus on the Skin Cancer dataset. The evaluation was conducted using the following metrics also summarized in Table 1: Average Accuracy, Total Execution Time, Precision (Cancer), Recall (Cancer), F1-Score (Cancer), Precision (Non-Cancer), Recall (Non-Cancer), F1-Score (Non-Cancer), and Overall Accuracy. The analysis clarifies the subtleties of various CNN architectures' effectiveness, particularly when using transfer learning strategies. Through an examination of individual outcomes for every model, this research offers valuable perspectives on the wider consequences of CNN architecture choice and transfer learning techniques in relation to the categorization of skin cancer.

4.1 AlexNet

Despite being a CNN architectural pioneer, AlexNet only showed an average accuracy of 66.78% over 30 iterations. There appears to be a significant imbalance between the cancer (35%) and non-cancer (89%) accuracy scores. With a recall of 67%, Cancer performed comparatively better. But the overall classification report and accuracy indicate that AlexNet might not be as reliable for this particular binary image classification task. It has the fastest execution with 333.09 seconds.

Table 1. Performance metrics and definitions.

Metric	Definition
Average Accuracy	The average percentage of skin cancer images correctly classified by each model
Total Execution Time	The time taken by each model to complete the classification task
Precision (Cancer)	The proportion of images correctly classified as cancerous out of all images the model classified as cancerous.
Recall (Cancer)	The proportion of cancerous images correctly identified by the model.
F1-Score (Cancer)	The harmonic mean of Precision (Cancer) and Recall (Cancer), providing a balanced measure of model performance for cancer classification.
Precision (Non-Cancer)	The proportion of images correctly classified as non-cancerous out of all images the model classified as non-cancerous.
Recall (Non-Cancer)	The proportion of non-cancerous images correctly identified by the model.
F1-Score (Non-Cancer)	The harmonic mean of Precision (Non-Cancer) and Recall (Non-Cancer), providing a balanced measure of model performance for non-cancer classification.
Overall Accuracy	The average accuracy of each model across classifying both cancerous and non-cancerous images.

4.2 EfficientNet

EfficientNet achieved an average accuracy of 69.36%, outperforming AlexNet. Better discriminative capacity was demonstrated by the precision values for both cancer (38%) and non-cancer (86%), which were more evenly distributed. Recall numbers (non-cancer: 78%, cancer: 50%) add to a performance that is generally more dependable. The reduced execution time of 512.05 seconds highlights the effectiveness of EfficientNet in managing activities related to binary image categorization.

4.3 MobileNetV3

With precision and recall numbers showing disparities between Cancer (precision: 25%, recall: 60%) and Non-Cancer (precision: 83%, recall: 53%), MobileNetV3 demonstrated an average accuracy of 59.97%. It's possible that MobileNetV3 isn't the best option for this particular binary image classification task because of its lower accuracy and differences in performance measures. To improve its performance, more research or adjustment might be required. It has an average execution time compared to others, with a total execution time of 399.00 seconds.

4.4 ResNet50

ResNet showed itself to be a strong performer, averaging 75.10% accuracy. A robust discriminative model was bolstered by the well-balanced accuracy values for Cancer (39%) and Non-Cancer (90%). The recall values (72%, 72%, and 69% for cancer and non-cancer, respectively) highlight Resnet's superior performance in binary image classification tasks. With a total execution time of 471.04 seconds, ResNet strikes a balance between efficiency and precision, making it a strong option.

4.5 VGG19

VGG19 showed competitive performance, averaging 72.63% accuracy. The precision values for Non-Cancer (95%) and Cancer (43%) show that the categorization strategy is balanced. The effectiveness of VGG19 was further supported by the well-distributed recall values for both classes. For real-time applications, the 828.54 second total execution time presents certain practical issues.

4.6 Overall Comparison

This study shows that CNN architectures for skin cancer classification trade off performance balance, accuracy, and speed. ResNet50 is the best at both accurately classifying skin cancer and striking a balance between precision and recall, as seen by its superior average accuracy (75.10%) and F1-Score (Cancer) (50%). Even though it loses some accuracy, MobileNetV3 has the quickest execution time (399.00s), which makes it perfect for applications where speed is of the essence, such as real-time screening. While VGG19 has the greatest F1-Score (Non-Cancer) (81%) and emphasizes non-cancer categorization, it is not as accurate overall as ResNet50. In the end, the best model selection depends on the requirements of the application. MobileNetV3 may be appropriate if real-time screening is the top priority. ResNet50 could be the ideal option for a conclusive diagnosis requiring a high degree of precision. VGG19 may be useful for applications that need to strike a compromise between accuracy and non-cancer categorization.

Table 2. Quantitative classification results of different CNN architectures.

	AlexNet	**EfficientNet**	**MobilNetV3**	**ResNet50**	**VGG19**
Mean Accuracy	66.78%	69.36%	59.97%	75.10%	72.63%
Total Runtime	333s	512s	399s	471s	823s
Precision (Cancer)	35%	38%	25%	39%	43%
Recall (Cancer)	67%	50%	60%	69%	86%
F1-Score (Cancer)	46%	43%	35%	50%	57%
Precision (Non-Cancer)	89%	86%	83%	90%	95%
Recall (Non-Cancer)	67%	78%	53%	72%	70%
F1-Score (Non-Cancer)	76%	82%	65%	80%	81%
Overall Accuracy	67%	73%	54%	71%	74%

5. Conclusion

To sum up, this research explored the field of binary image classification for the purpose of identifying skin cancer, with a particular emphasis on comparing different CNN architectures and making strategic use of transfer learning. The study employed well-known CNN models to the Skin Cancer Binary Classification Dataset, such as ResNet50, MobileNetV3, VGG19, AlexNet and EfficientNet.

The results show that the models' performance in binary image classification tasks is highly influenced by the CNN architecture selected. Top performances EfficientNet and ResNet50 demonstrated a fair trade-off between execution time and accuracy. However, MobileNetV3 performed worse, highlighting how crucial it is to choose the right architecture for a given task.

The results highlight how transfer learning might improve CNN's performance in diagnosing skin cancer, especially in situations when datasets are few.

For future studies, scientists may try optimizing CNN architectures to perform better on particular subtypes of skin cancer or look at how well models generalize to a variety of dermatological datasets. Furthermore, improvements in transfer learning methodologies and the investigation of new CNN architectures may enhance the precision and effectiveness of automated systems for the detection of skin cancer.

This study makes a significant contribution to the area of medical image analysis by directing the development of automated techniques for detecting skin cancer and encouraging further advancements in diagnostic precision.

Authors' Contributions

All authors contributed equally to the study.

Statement of Conflicts of Interest

There is no conflict of interest between the authors.

Statement of Research and Publication Ethics

The author declares that this study complies with Research and Publication Ethics.

References

Al-Mekhlafi, Z.G., Senan, E.M., Rassem, T.H., Mohammed, B.A., Makbol, N.M. et al. (2022). Deep learning and machine learning for early detection of stroke and haemorrhage. *Computers, Materials & Continua*, 72(1), 775-796. https://doi.org/10.32604/cmc.2022.024492

Al-saedi, D. K. A., & Savaş, S. (2022). Classification of Skin Cancer with Deep Transfer Learning Method. *IDAP-2022 : International Artificial Intelligence and Data Processing Symposium*, 202-210. https://doi.org/10.53070/bbd.1172782

Bai, X., Shang, X., Zhong, S., Hu, P., & Chang, C. (2019, May). Iterative constrained energy minimization convolutional neural network for hyperspectral image classification. *Algorithms, Technologies, and Applications for Multispectral and Hyperspectral Imagery XXV*. https://doi.org/10.1117/12.2519046

Bu, Y., Jiang, X., Tian, J., Hu, X., Han, L., Huang, D., … & Luo, H. (2022). Rapid argnondestructive detecting of sorghum varieties based on hyperspectral imaging and convolutional neural network. *Journal of the Science of Food and Agriculture*, 103(8), 3970-3983. https://doi.org/10.1002/jsfa.12344

Craythorne, E., & Al-Niami, F. (2017). Skin tumours: Skin cancer. Medicine, 45(7), 431-434. https://doi.org/10.1016/j.mpmed.2017.04.003

Deng, T. and Wu, Y. (2022). Simultaneous vehicle and lane detection via mobilenetv3 in car following scene. *Plos One*, 17(3), e0264551. https://doi.org/10.1371/journal.pone.0264551

Dildar, M., Akram, S., Irfan, M., Khan, H. U., Ramzan, M., Mahmood, A. R., … & Mahnashi, M. H. (2021). Skin cancer detection: a review using deep learning techniques. *International Journal of Environmental Research and Public Health*, 18(10), 5479. https://doi.org/10.3390/ijerph18105479

Du, X., Si, L., Li, P., & Yun, Z. (2023). A method for detecting the quality of cotton seeds based on an improved resnet50 model. *Plos One*, 18(2), e0273057. https://doi.org/10.1371/journal.pone.0273057

Garg, S., Mohanty, D. P., Thota, S. P., & Moharana, S. (2021). A simple approach to image tilt correction with self-attention mobilenet for smartphones. *In Proceedings of the British Machine Vision Conference 2021* (pp. 1-12). https://doi.org/10.48550/arxiv.2111.00398

Geng, Y., Huang, J., Yang, J., & Zhang, S. (2022). Spectrum sensing for cognitive radio based on feature extraction and deep learning. *Journal of Physics Conference Series*, 2261(1), 012016. https://doi.org/10.1088/1742-6596/2261/1/012016

Guo-zhan, W., Feng, A., Gu, C., & Liu, X. (2023). Yolo-dfd: a lightweight method for dog feces detection based on improved yolov4. *Journal of Sensors*, 2023, 1-11. https://doi.org/10.1155/2023/5602595

Hasan, M.K., & Aleef, T.A. (2019). Automatic Mass Detection in Breast Using Deep Convolutional Neural Network and SVM Classifier. https://doi.org/10.48550/arxiv.1907.04424

Hu, J., Qi, Y., & Wang, J. (2022). Skin disease classification using mobilenet-rsesk network. *Journal of Physics: Conference Series*, 2405(1), 012017. https://doi.org/10.1088/1742-6596/2405/1/012017

Huang, Y., Vadloori, S., Chu, H., Kang, E., Wu, W., & Fukushima, Y. (2020). Deep learning models for automated diagnosis of retinopathy of prematurity in preterm infants. *Electronics*, 9(9), 1444. https://doi.org/10.3390/electronics9091444

Islam, M. S., & Panta, S. (2024). Skin cancer images classification using transfer learning techniques. *arXiv*.

Jia, X., & Li, D. (2022). TFCN: Temporal-Frequential Convolutional Network for Single-Channel Speech Enhancement. https://doi.org/10.48550/arxiv.2201.00480

Kamiri, J., Wambugu, G. M., & Oirere, A. M. (2022). A comparative study of deep learning and transfer learning in detection of diabetic retinopathy. International *Journal of Computer Applications Technology and Research*, 11(07), 247-254. https://doi.org/10.7753/ijcatr1107.1001

Kashani, H. B., Jodeiri, A., Goodarzi, M. M., & Rezaei, I. S. (2019). Speech enhancement via deep spectrum image translation network. *In Proceedings of the 2019 26th National and 4th International Iranian Conference on Biomedical Engineering (ICBME)* (pp. 1-7). IEEE. https://doi.org/10.1109/ICBME49163.2019.9030421

Laschowski, B., McNally, W., Wong, A., & McPhee, J. (2021). Computer vision and deep learning for environment-adaptive control of robotic lower-limb exoskeletons. In Proceedings of the Annual International *Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 4631-4635). IEEE. https://doi.org/10.1109/EMBC46164.2021.9630064

Li, Y., Zheng, H., Huang, X., Chang, J., Hou, D., & Lu, H. (2022). Research on lung nodule recognition algorithm based on deep feature fusion and mkl-svm-ipso. *Scientific Reports*, 12(1). https://doi.org/10.1038/s41598-022-22442-3

Liu, Y., Tang, K., Cai, W., Chen, A., Zhou, G., Li, L., … & Liu, R. (2022). MPC-STANet: Alzheimer's disease recognition method based on multiple phantom convolution and spatial transformation attention mechanism. *Frontiers in Aging Neuroscience*, 14. https://doi.org/10.3389/fnagi.2022.918462

Mvoulana, A., Kachouri, R., & Akil, M. (2021). Fine-tuning convolutional neural networks: A comprehensive guide and benchmark analysis for glaucoma screening. *In Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR)* (pp. 4677-4684). IEEE. https://doi.org/10.1109/icpr48806.2021.9412199

Prima, B. and Bouhorma, M. (2020). Using transfer learning for malware classification. The International Archives of the Photogrammetry *Remote Sensing and Spatial Information Sciences*, XLIV-4/W3-2020, 343-349. https://doi.org/10.5194/isprs-archives-xliv-4-w3-2020-343-2020

Rashid, J., Ishfaq, M., Ali, G., Saeed, M. R., Hussain, M., Alkhalifah, T., Alturise, F., & Samand, N. (2022). Skin cancer disease detection using transfer learning technique. *Applied Sciences*, 12(11), 5714. https://doi.org/10.3390/app12115714

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., … & Li, F. (2015). ImageNet large scale visual recognition challenge. International *Journal of Computer Vision*, 115(3), 211-252. https://doi.org/10.1007/s11263-015-0816-y

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. (2018). Mobilenetv2: inverted residuals and linear bottlenecks. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. https://doi.org/10.1109/cvpr.2018.00474

Singh, B. and Sharma, D. K. (2021). Predicting image credibility in fake news over social media using multi-modal approach. *Neural Computing and Applications*, 34(24), 21503-21517. https://doi.org/10.1007/s00521-021-06086-4

Sobczak, S., & Kapela, R. (2022). Hybrid restricted Boltzmann machine–convolutional neural network model for image recognition. *IEEE Access*, 10, 24985-24994. https://doi.org/10.1109/access.2022.3155873

Suciu, O., Coull, S. E., & Johns, J. (2019). Exploring adversarial examples in malware detection. *2019 IEEE Security and Privacy Workshops (SPW*), 8-14. https://doi.org/10.1109/spw.2019.00015

Tan, M. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. *International Conference on Machine Learning (ICML)*. https://doi.org/10.48550/arxiv.1905.11946

Tinnathi, S., & Sudhavani, G. (2022). Copy-move forgery detection using superpixel clustering algorithm and enhanced GWO-based AlexNet model. *Cybernetics and Information Technologies*, 22(4), 91-110. https://doi.org/10.2478/cait-2022-0041

Tufail, A., Ma, Y., Kaabar, M., Rehman, A., Khan, R., & Cheikhrouhou, O. (2021). Classification of initial stages of Alzheimer's disease through PET neuroimaging modality and deep learning: Quantifying the impact of image filtering approaches. *Mathematics*, 9(23), 3101. https://doi.org/10.3390/math9233101

Ullah, I., & Mahmoud, Q. (2021). Design and development of a deep learning-based model for anomaly detection in IoT networks. *IEEE Access*, 9, 103906-103926. https://doi.org/10.1109/ACCESS.2021.3094024

Wang, Y., Chun, X., Zhu, B., Wang, M., Wang, T., Ni, P., …, & Hu, J. (2022). A new non-invasive tagging method for leopard coral grouper (Plectropomus leopardus) using deep convolutional neural networks with PDE-based image decomposition. *Frontiers in Marine Science*, 9, 1093623. https://doi.org/10.3389/fmars.2022.1093623

Xu, Y., Zhao, B., Zhai, Y., Chen, Q., & Zhou, Y. (2021). Maize diseases identification method based on multi-scale convolutional global pooling neural network. *IEEE Access*, 9, 27959–27970. https://doi.org/10.1109/access.2021.3058267