Skin Cancer Detection: Using Deep Learning and Transfer Learning Techniques

Rami Djekoun*, Nadir Farah
LABGED Laboratory, Department of Computing, Badji Mokhtar University Annaba, Algeria
ramidjekoun23@gmail.com , farah@labged.net

* *Corresponding author: ramidjekoun23@gmail.com*

Abstract: Skin cancer is one of the most perilous forms of cancer, stemming from unrepaired DNA damage in skin cells, leading to genetic abnormalities or mutations. Its tendency to slowly spread to other body parts underscores the critical importance of early detection. Researchers have thus devised various early detection methods, utilizing parameters such as symmetry, color, size, and shape of lesions. An innovative approach employing deep learning and transfer learning has emerged, achieving up to a 95% correct classification rate of malignant lesions from skin images. This breakthrough offers hope in the fight against melanoma by enabling earlier and more precise diagnoses, crucial for swift treatment. However, the scarcity of skilled dermatologists globally remains a challenge in addressing current healthcare needs. This article sheds light on the challenges and clinical testimonies surrounding this major advancement in skin cancer treatment, illustrating both the benefits and hurdles of integrating AI techniques in dermatology and medicine.

Keywords: Deep Learning, Transfer Learning, VGG16, VGG19, Res-net50, Melanoma, Skin Lesion.

1 INTRODUCTION

The number of cancer patients is increasing due to smoking, environmental changes, different types of radiation, viruses, alcohol, diet, and lifestyle [1]. Among the myriad forms of cancer, skin cancer stands out as one of the most prevalent and hazardous. Manifesting as abnormal growths of skin cells, skin cancer is spreading worldwide and is a perilous disease [2]. The overwhelming increase in its incidence rates, particularly of melanoma, has grown over 300% from 1990 to 2018 just in the US [3]. The recorded new skin cancer cases rate in the USA is around 5.4 million a year [4]. Notably, melanoma, a particularly lethal form of skin cancer, has witnessed a 53% rise in yearly diagnoses according to the World Health Organization (WHO), with its mortality rate projected to surge over the next decade [3] [4]. More than two people die from skin cancer in the United States every hour. Prior to the 1980s, melanoma detection was performed by observing macroscopic features, as they were usually detected when their size was large. This made early detection unlikely, and death rates continued to rise. In 1985, a research team at New York University developed the term ABCD, which stands for Asymmetry, Border irregularity, Color variegation, and Diameter (ABCD), as a simple and effective tool to educate the public about the early identification of melanoma. After 1990, screening by physicians and the use of basic full-body imaging became the common methods for detecting melanoma at an early stage. Over time, computer-augmented digital analysis became the new phenomenon due to its high sensitivity and specificity of detection compared to manual dermoscopy. It is difficult for a dermatologist to detect skin cancer from a dermoscopy image of a skin lesion [5]. Basal cell carcinoma (BCC) and squamous cell carcinoma (SCC) are by far the most prevalent forms of skin cancer [6]. With the advancement of healthcare and medicine, computer diagnostic solutions assisted by artificial intelligence (AI) are making significant contributions, particularly in the field of medical imaging. Recent advances in deep learning, particularly convolutional neural networks (CNNs), have shown promising results in the automatic detection and classification of skin lesions. In 2018, Tschandl et al. published HAM10000. Using large databases such as HAM10000, which includes 10,000 dermoscopy

C. A. Kerrache et al. (eds.), *Proceedings of the International Conference on Emerging Intelligent Systems for Sustainable Development (ICEIS 2024)*, Advances in Intelligent Systems Research 184,
https://doi.org/10.2991/978-94-6463-496-9_20

images with confirmed diagnoses, researchers have trained deep learning models to detect malignant lesions. This progression has the potential to improve early detection rates, reduce unnecessary biopsies, and ultimately save lives [7].

2 CNN EVOLUTION IN DERMOSCOPY IMAGES

Advances in dermoscopic image analysis have been propelled by the utilization of convolutional neural networks (CNNs). Codella et al. (2017) [8] were among the early contributors to this field, employing the Caffe architecture as a feature extractor. Esteva et al. (2017) [6] demonstrated, for the first time, the comparable classification capability of CNNs to expert dermatologists. Leveraging a vast dataset of over 129,000 dermoscopic images, they showcased the robustness and generalizability of CNN methods. Hansler et al [9]. (2018) delved deeper into dermoscopic melanoma identification, comparing the performances of a CNN and 58 dermatologists. They illustrated the effectiveness of CNNs in identifying melanoma, highlighting their significant clinical diagnostic potential. Tschandl et al [7]. (2019) developed a method to detect melanoma early by identifying specific melanoma signatures from dermoscopic images. Their analysis underscored CNNs' high sensitivity in detecting malignant melanocytic lesions, providing a valuable tool for early melanoma detection. Menegola et al [10]. (2020) conducted a comprehensive analysis, evaluating the progress made and the remaining challenges in transitioning the use of CNNs into clinical dermatology applications. Finally, Fekry Olayah et al [11]. (2023) conducted research on the early detection of skin lesions using AI techniques for dermoscopic image analysis, focusing on integrating CNN features for improved accuracy in identifying skin abnormalities. Their study aimed to contribute to the early detection of skin cancer through advanced image analysis methods. This collective effort clearly delineates the impact that CNNs have had on the analysis of dermoscopic images, promising improved detection and treatment of patients with skin lesions in the future.

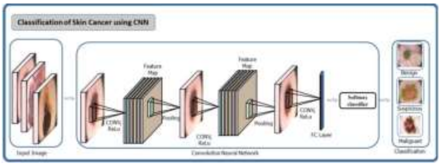

Fig 1. CNN feature extraction framework for skin cancer. Adapted from Saba T [12]. "Recent advancement in cancer detection using machine learning.

3 DEEP LEARNING AND TRANSFER LEARNING

3.1 Convolutional Neural Network for Image Classification

Artificial neural networks (ANN) are composed of neurons inspired by the biological neurons in our brain. Among them, convolutional neural networks (CNN) represent a major advancement, often used for image classification tasks due to their ability to recognize specific objects even in different appearances, thanks to their understanding of the invariance of translation [13]. Unlike feedforward neural networks, which often fail to recognize objects when they are not centred in the image, CNNs can identify objects regardless of their position or orientation in the image [14]. This capability is made possible by the structure of CNNs, including convolution and pooling layers, which make it possible to detect patterns in images and reduce data complexity [15]. In practice, CNNs have significant advantages, including their ability to be trained on raw, unprocessed datasets, thereby reducing the time and resources required for training models [16]. The CNN's phases include several key steps: convolution, where filters slide over the image to detect local features; pooling, which reduces the dimensionality of feature maps while retaining the most important information. Fully connected layers, which combine local features to perform global classifications; and finally, the output phase, which uses activation functions like SoftMax to produce probabilities of each object class. This architecture allows CNNs to perform robust and efficient

3.2 Transfer learning

Transfer learning is a method of learning where a model learns about one problem before this serves as the starting point for another task. This is a suitable approach for problems when a procedure near the primary issue already exists and the related task requires many data [17]. Transfer learning uses the technique of feature extraction from a pre-trained model; this eliminates the need for developers to start over when training a model. A TL model is typically trained on a large dataset (for example, ImageNet) [18] and the related parameters obtained from the trained model can then be used with a custom neural network for any other related application. These types of models can be used directly for predictions on new tasks or in any other related application training processes of the model. In this study focusing on transfer learning, four widely utilized pre-trained neural network models in computer vision are employed: VGG16 and VGG19, introduced by Simonyan and Zisserman [19], feature a deep structure comprising convolution layers followed by pooling layers. ResNet50, proposed by He et al. (2016), utilizes "deep residual learning" to mitigate performance degradation associated with increasing network depth through residual connections [20]. InceptionV3, a pivotal model in computer vision, was devised by Szegedy et al. [21], featuring a novel architecture utilizing Inception modules for feature extraction across varying spatial scales [21].

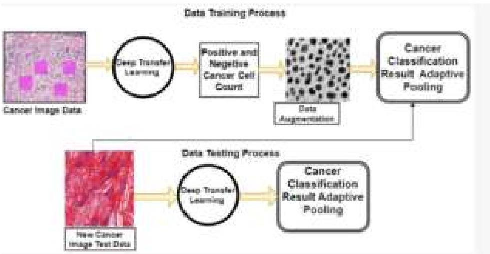

Fig 2. Classification of dermoscopic images with Transfer learning [22]

4 Materials and Methods

4.1 Description of HAM10000 Dataset

In recent years, various corpora of skin pathologies made up of dermoscopy images have enabled the implementation of CAD systems for the analysis of skin cancer [16]. The HAM10000 dataset, which includes 10015 dermatoscopic images include a representative selection of all major diagnostic categories in the field of pigmentary lesions: actinic and intraepithelial (akiec), basal cancer cells (bcc), benign keratinosis-like lesions (solar/seborrheic lentigines and lichen planus similar, bkl), dermatofibroma (df), melanoma (mel), melanocytic nevi (nv) and vascular lesions (angiomas, angiokeratomas, pyogenic granulomas and hemorrhage, vasc). More than 50% of lesions are confirmed by histopathology (histo), while for the other cases the fundamental truth is obtained using in- depth evaluation (deep evaluation), expert consensus (consensus); in vivo confocal microscopy (confocal microscopy). The file contains samples with multiple images. Which can be tracked by the lesion_id column in the HAM10000_metadata file. The figure depicts the classification of images within the HAM10000 Dataset into various diagnostic categories, providing an overview of the distribution of skin pathologies represented in the dataset.

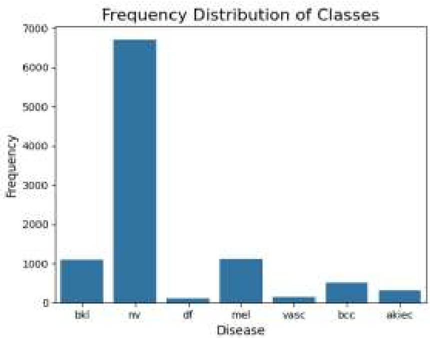

Fig3. Distribution of Classes in the HAM10000

4.2 Pre-processing

In this study, we present a comprehensive preprocessing pipeline tailored for dermatological image classification tasks utilizing Convolutional Neural Networks (CNNs) [21]. Our preprocessing pipeline involves various steps including data loading, data cleaning, augmentation, and normalization to prepare the dataset for efficient training and evaluation of CNN models. Architecture of the preprocessing pipeline in dermatologic image classification for CNNs.

4.3 Data Cleaning

It is the procedure to load the dermatologic image dataset into processing. We will also fetch the related metadata along with the image. The dataset is present in a single directory but divided already into two directories for convenience of this work. We will iterate through each file in the directory and get a list of all the files present in each directory. The metadata is in the CSV format and is present at a certain location. The CSV file is used by the read_csv method of the pandas dataframe.

4.4 Class Imbalance Handling

In most of the medical image classification models, class imbalance is a big issue since the number of images of a particular class would be inferior to the others. During the process of model training, this will hinder the model to get trained well leading to lesser performance. Therefore, we will use the Random over Sampler technique from the imbalanced-learn library to balance the classes. It creates synthetic samples for the minority class to equalize the number of samples for the majority as well as the minority classes.

4.5 Data Augmentation and Normalization

It is the most important step in the preprocessing of the image. Training a deep learning model without doing some data augmentation cannot be considered. The data is artificially increasing the size of the dataset by using different methods of data augmentation. We will apply the ImageDataGenerator method from the TensorFlow library to apply real-time data augmentation on the fly while training the model. The rescaling of the images is done so that the pixel values remain between 0 and 1.

4.6 Random Over Sampler Technique

The Random over Sampler technique is a method commonly used in machine learning, particularly in addressing classification problems with imbalanced class distributions. The primary objective of Random over Sampling is to mitigate the imbalance by artificially increasing the number of samples in the minority class within the training dataset. This is achieved by randomly duplicating instances from the minority class, thereby ensuring a more balanced representation across all classes. By doing so, the potential bias introduced by class imbalance is reduced, facilitating improved generalization and predictive performance of machine learning models [23].

4.7 Image Data Generator Method

The Image Data Generator method is a widely adopted technique for data augmentation during the training of convolutional neural network (CNN) models, typically used in image classification tasks. Rather than training on a static dataset, the Image Data Generator generates augmented versions of original images dynamically during each epoch of training. Augmentations may include random rotations, zooms, horizontal or vertical shifts, flips, and other transformations. This approach enriches the training set, enabling the model to generalize better and mitigate overfitting by exposing it to a wider variety of image variations and scenarios [25, 26].

4.8 Network Training

In order to train the network, various pre-trained architectures are used, MobileNetV2, VGG16, VGG19, InceptionV3 and a CNN. The custom CNN consists of multiple convolution layers followed by max pooling layers and fully connected layers. The transfer learning models were loaded with their pre-trained weights on the ImageNet dataset and were adapted for the skin lesion classification task. Each model was trained with the Adam optimizer and a categorical cross-entropy loss function with a batch of 128 elements. The training set was divided into training and evaluation sets, and models were evaluated on the training set to track their performance across epochs. Regularization methods such as dropout and batch normalization have been used to limit over fitting.

5 RESULTS

The classification performance of the CNN model was found to be highly impressive, with an accuracy rate of 99.04% and validation accuracy of 96.67%. The recorded losses were exceptionally low, with values of 0.0297 for regular loss and 0.1698 for validation loss. These results clearly demonstrate the efficacy of employing CNNs for the precise identification of different types of skin lesions in dermoscopic images. In a similar vein, the performance of the MobilenetV2 architecture was also observed to be robust, achieving a commendable accuracy rate of 97.44% and a validation accuracy of 98.08%. The corresponding loss values were recorded at 0.0766 and 0.0740, respectively. These outcomes indicate the potential utility of utilizing more lightweight neural network structures for the efficient classification of dermoscopic images. Conversely, the VGG16, VGG19, and InceptionV3 models delivered competitive performance outcomes during our experiments; however, their accuracy rates ranged between 91.65% to 93.55%, with validation accuracies falling within 95.12% to 95.94%. These results suggest marginally lower performance levels compared to the CNN and MobilenetV2 architectures. Nevertheless, these models still displayed strong classification capabilities for accurately identifying skin lesions. The results agree with previous researches that have investigated CNNs for analyzing dermoscopic images. For example, Esteva et al. (2017) concluded that the CNNs performed almost as well as those of the skin doctors, who are professionals in this field, which is similar to this work that has high accuracy. Similarly, Tschandl et al. (2019) say that using CNNs also highlights their sensitivity in detecting malignant lesions. This supports the findings with respect to strong classification ability. Moreover, such lightweight architectures as MobileNetV2 are being explored to develop efficient models for clinical applications (Menegola et al., 2020). Consequently, our findings add up to the growing evidence base supporting the use of CNN in dermatoscopy, which holds great promise for enhancing diagnostic accuracy and delivery of care among dermatologists.

Table 1. Results of Different Models in Skin Cancer Classification.

Models	Training loss	Training Accuracy	Validation loss	Validation accuracy
CNN	0.02	0.99	0.16	0.96
MobilenetV2	0.07	0.97	0.07	0.98
VGG16	0.21	0.92	0.13	0.95
VGG19	0.24	0.91	0.14	0.95
InceptionV3	0.21	0.93	0.15	0.95

This section describes our findings, along with three figures illustrating the performance of various skin cancer classification models. Figure 4 shows a comparison of training accuracy between models and provides insight into how effectively they learn from training data. Meanwhile, Figure 5 shows a comparison of training losses between models and highlights the differences in their ability to minimize errors during training. Finally, Figure 6 provides a graphical representation of the best validation scores obtained by the models and illustrates their relative performance when validating against different data in the training set.

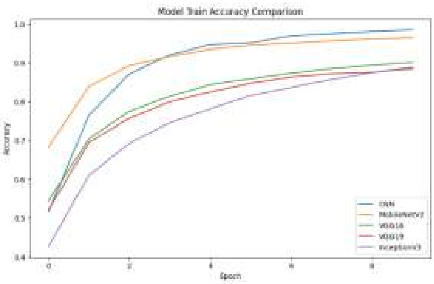

Fig 4. Training Accuracy Models Comparing

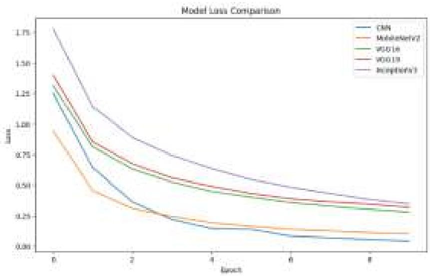

Fig5. Loss Models Comparing

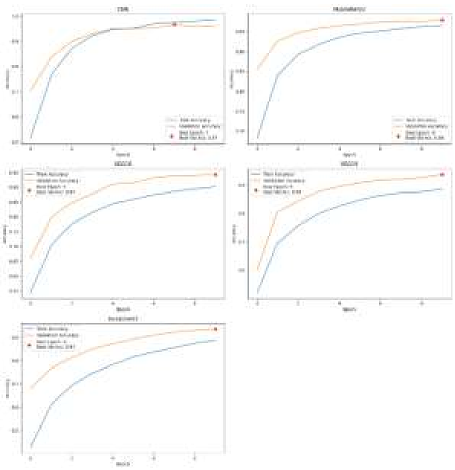

Fig 6. Graphic Representation of Best Validation Values of The Models

6 CONCLUSIONS

In sum, this study shows how well convolutional neural networks (CNNs) work for sorting dermoscopic pictures. They are very good at identifying different skin problems with a lot of accuracy. The findings show that CNN and MobileNetV2 models do really well, with accuracies of more than 99% and 97%. Even though VGG16, VGG19, and InceptionV3 models are a bit lower, they are still good at sorting skin problems right. The obtained results match other studies, showing that CNNs are useful in dermoscopic picture review. They work almost as well as expert skin doctors, as Esteva et al. (2017) showed [6], and can detect bad lesions, like Tschandl et al. (2019) emphasized [7]. In addition, looking at lightweight models like MobileNetV2 supports efforts to make good models for clinical jobs, as Menegola et al. (2020) pointed out [10]. The presented work adds to the proof that CNNs work well for analysing dermoscopic pictures. The presented models do a good job and could help make diagnoses better and take care of patients better in skin care. As we learn more about deep learning, more studies like these will help the use of computer smarts in skin care even more.

REFERENCES

1. Alam, T.M., Khan, M.M.A., Iqbal, M.A. et al. Cervical cancer prediction through different screening methods using data mining. Int. J. Adv. Comput. Sci. Appl. 10, 388–396 (2019). https://doi.org/10.14569/IJACSA.2019.0100547

2. Siegel, R.L., Miller, K.D., Jemal, A. Cancer statistics, 2018. CA Cancer J Clin 68, 7–30 (2018). https://doi.org/10.3322/caac.214.

3. Ferlay, J., Colombet, M., Soerjomataram, I. et al. Cancer statistics for the year 2020: An overview. Int J Cancer 149, 778–789 (2021). https://doi.org/10.1002/ijc.33867

4. Brunssen, A., Waldmann, A., Eisemann, N., Katalinic, A. Impact of skin cancer screening and secondary prevention campaigns on skin cancer incidence and mortality: A systematic review. J Am Acad Dermatol 76, 129–139.e10 (2017). https://doi.org/10.1016/j.jaad.2016.08.057

5. Niino, M., Matsuda, T. Age-specific skin cancer incidence rate in the world. Jpn J Clin Oncol 51, 848–849 (2021). https://doi.org/10.1093/jjco/hyab103

6. Esteva, A., Kuprel, B., Novoa, R.A. et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 542, 115–118 (2017). https://doi.org/10.1038/nature21056.

7. Tschandl, P., Rosendahl, C., Kittler, H. The HAM10000 dataset, a large collection of multisource dermatoscopic images of common pigmented skin lesions. Sci Data 5, 180161 (2018). https://doi.org/10.1038/sdata.2018.161

8. Codella, N.C.F., Cai, J., Abedini, M. et al. Deeplearning, sparse coding, and SVM for melanoma recognition in dermoscopy images. In: MLMI 2015, pp. 118–126 (2015). https://doi.org/10.1007/978-3-319-25548-1_14

9. Haenssle, H.A., Fink, C., Schneiderbauer, R. et al. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Ann Oncol 29, 1836–1842 (2018). https://doi.org/10.1093/annonc/mdy166

10. Menegola, A., Tavares, J., Fornaciali, M. Convolutional neural networks for skin lesion classification in dermatology: a review. J Ambient Intell Human Comput 1–19 (2020). https://doi.org/10.1007/s12652-020-02002-1

11. Olayah, F., Senan, E.M., Ahmed, I.A., Awaji, B. AI Techniques of Dermoscopy Image Analysis for the Early Detection of Skin Lesions Based on Combined CNN Features. Diagnostics 13, 1314 (2023). https://doi.org/10.3390/diagnostics13071314

12. Saba, T. Recent advancement in cancer detection using machine learning: Systematic survey of decades, comparisons and challenges. J Infect Public Health (2020). https://doi.org/10.1016/j.jiph.2020.10.010

13. Krizhevsky, A., Sutskever, I., Hinton, G.E. Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, 25 (2012).

14. Simonyan, K., Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014).

15. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... Rabinovich, A. Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2015). https://doi.org/10.1109/CVPR.2015.7298594

16. He, K., Zhang, X., Ren, S., Sun, J. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2016). https://doi.org/10.1109/CVPR.2016.90

17. Sarraf, S., Tofighi, G. Classification of Alzheimer's Disease Using fMRI Data and Deep Learning Convolutional Neural Networks. arXiv preprint arXiv:1603.08631v1 [cs.CV] (2016).

18. He, K., Zhang, X., Ren, S., Sun, J. Deep residual learning for image recognition. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778 (2016). https://doi.org/10.1109/CVPR.2016.90

19. Simonyan, K., Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv preprint arXiv:1409.1556 (2014).

20. He, K., et al. Deep Residual Learning for Image Recognition. In: Proceedings of the IEEE conference on computer vision and pattern Recognition (CVPR) (2016). (CVPR).

21. Szegedy, C., et al. Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR) (2016).

22. Bhuiyan, M.R., Abdullah, J. Detection on Cell Cancer Using the Deep Transfer Learning and Histogram Based Image Focus Quality Assessment. Sensors 22(18), p.7007 (2022). https://doi.org/10.3390/s22187007

23. Buda, M., Maki, A., Mazurowski, M.A. A systematic study of the class imbalance problem in convolutional neural networks. Neural Netw 106, 249–259 (2018). https://doi.org/10.1016/j.neunet.2018.07.011

24. Shorten, C., Khoshgoftaar, T.M. A survey on Image Data Augmentation for Deep Learning. J Big Data 6, 60 (2019). https://doi.org/10.1186/s40537-019-0197-0

25. Perez, L., Wang, J. The effectiveness of data augmentation in image classification using deep learning. Convolutional Neural Networks Vis Comput 6, 11–25 (2017). https://doi.org/10.1145/3195431