Classification of Skin Cancer Using ResNet and VGG Deep Learning Network

Clara Lavita Angelina [1,2], Riki Umami Sanaz Ulfitria[2]

{d11010212@gemail.yuntech.edu.tw [1], m10812096@gemail.yuntech.edu.tw [2]}

Graduate School of Engineering Science and Technology, National Yunlin University of Science and Technology, Yunlin, Taiwan [1],
Department of Electrical Engineering, National Yunlin University of Science and Technology, Yunlin, Taiwan [2]

Abstract. Skin cancer develops as a result of the unregulated growth of mutations in DNA caused by a variety of factors. Accurate classification of skin cancer is crucial to distinguishing high malignant potential from benign potential. In this paper, we utilized the ResNet-50, ResNet-101, VGG16, and VGG19 networks to develop a deep learning model for the classification of skin cancer. A total of 3,217 images from two types of skin cancer are used as a training dataset. Our method can automatically classify the feature type of skin cancer and record the prediction result in each image. The method has been evaluated on 80 test images. The experimental result shows that the ResNet50, ResNet101, VGG16, and VGG19 achieve accuracy up to 78.75%, 75%, 83.75%, and 73.75% respectively.

Keywords: Skin Cancer; Classification; Deep Learning; ResNet; VGG; Benign; Malignant

1. Introduction

Skin cancer is a prevalent kind of cancer characterised by the unregulated proliferation of cells within the skin. The development of skin cancer can be attributed to the exposure of skin cells to ultraviolet (UV) radiation emitted by the sun or artificial tanning devices, leading to the proliferation of malignant tumors. As to the World Health Organisation (WHO), the incidence of skin cancer accounts for approximately one-third of all cancer occurrences globally. Moreover, there is a consistent annual rise in the number of reported cases, attributed to the escalating penetration of solar ultraviolet (UV) radiation through the Earth's atmosphere [1]. Skin cancer develops as a result of the uneven growth of melanocytic skin cells.

Invasive and innocuous skin cancers are considered to be the most life-threatening. A benign neoplasm has the potential to develop but lacks the ability to metastasize. The identification of typical indications and symptoms of potentially cancerous benign skin growths, together with seeking medical attention for suspicious skin growths, is of utmost importance. Seborrheic keratosis, cherry angiomas, dermatofibromas, acrochordons (also known as skin tags), pyrogenic granulomas, and epidermal inclusion cysts represent a range of benign skin growths

ABEC 2023, September 21, Bengkalis, Indonesia
Copyright © 2024 EAI
DOI 10.4108/eai.21-9-2023.2342881

[2]. A hazardous cancerous is a type of neoplasm characterized by its cancerous nature, exhibiting the ability to metastasize and proliferate within the host's body. These entities possess the ability to infiltrate and colonize many bodily tissues and organs, proliferating and disseminating in an unregulated manner. Numerous malignant cutaneous neoplasms exhibit symptoms that can be discerned as preliminary indicators. A precursor refers to a cluster of atypical cells that possess the capacity to develop into malignant tumors. A precursor is another word for cancer. Certain types of skin growths that are considered precancerous exhibit varying levels of risk for developing cancer. Malignant skin growths encompass various types, such as melanoma, carcinoma, sarcoma, squamous cell carcinoma, and skin lymphoma[3].

The early identification of carcinoma of the skin has the potential to significantly reduce mortality rates. The workaround was offered by machine learning (ML) models. The application of deep learning, specifically the Convolutional Neural Network (CNN), has the potential to efficiently and cost-effectively facilitate the diagnosis of skin cancer through the categorization of images. The aforementioned resource has emerged as a crucial means of support for individuals experiencing extreme poverty. When considering the diagnosis of skin cancer by image classification, it can be observed that these machine-learning models exhibit higher levels of accuracy and efficiency. In today's environment, medical research is evolving. In the past, the identification of cancer of the skin was performed through manual means, resulting in a significant expenditure of time and resources [2]. However, due to recent advancements in advanced machine learning within the field of medical science, the process has been simplified. The study's methodologies propose the utilization of Convolutional Neural Networks (CNN) for malignancies of skin detection.

Due to its considerable potential for screening and identifying cancer in its early stages using dermoscopic pictures, hospitals have begun adopting image-based computer-aided diagnosis (CAD) technique. From an image processing standpoint, CAD systems examine numerous aspects like texture, color, and form to decide whether the data is malignant or benign. The identification of the specific type of skin cancer is crucial in cases where the disease is benign, as it facilitates the selection of the most suitable treatment approach. Since there are several forms of cancer, this categorization is very non-linear, and as a result, deep convolutional neural networks (DCNNs) can properly categorize them [4]. Numerous scholars have employed deep convolutional neural network (DCNN) architectures in their investigations on skin cancer datasets, aiming to enhance the efficacy of skin cancer detection methodologies.

This research aimed to benchmark four popular convolutional neural network architectures, namely ResNet-50, ResNet-101, VGG16, and VGG19, with no data preprocessing for the task of skin cancer classification. Skin cancer classification is a critical area in dermatology, and the choice of model architecture can significantly impact diagnostic accuracy.

The subsequent parts of the paper are organized in a particular manner: The second part presents a comprehensive survey of the extant scholarly works in the respective academic domain. The third section provides a comprehensive overview of the methodology and training procedure employed in the proposed deep convolutional neural network (DCNN). Additionally, it includes a comparative analysis between the DCNN and two other well-known architectures, namely VGG and ResNet. the fourth section summarises the findings and outcomes of the investigation. The final section summarises the main points and makes some potential suggestions for future research.

2. Related Works

Skin cancer has been identified using neural networks and training datasets. In the past, a detailed and comprehensive review of deep learning algorithms for the early detection of skin cancer was completed. It is being studied how a learning model is trained using pre-processed and trained datasets. The model is trained to understand the image's lesion features, such as symmetry, color, size, shape, and others, which are then used to diagnose skin cancer and distinguish benign from malignant skin cancer. Based on their work methods, this study compares multiple Convolutional Neural Network (CNN) models. Convolutional Neural Networks (CNNs) are widely employed in the field of computer vision for the purpose of picture recognition and categorization. [5].

S. Kalouche [6] and Dipu Chandra Malo et al. [7] have proposed CNN and VGG16 methods for skin cancer identification. The acquisition of features of afflicted skin cells involves dividing dermoscopic pictures using attribute extraction methodology. The characteristics of the affected skin cells are obtained following the segmentation of dermoscopic images by the utilization of the attribute extraction technique. They achieved an accuracy of 78% [6]. The characteristics of the affected skin cells are obtained after the segmentation of dermoscopic pictures using the feature extraction technique. They achieved an accuracy of 87.6% [7]. Nour Abuared et al. [8] found that VGG19-based CNN and transfer learning were effective techniques for assisting in skin cancer detection. The network's overall accuracy and loss suggest a decent result that may be improved further. The training technique is then discussed, tested, and assessed by measuring the total accuracy and loss of the network.

Ahmet Demr et al. [9] The objective of this project is to create a proficient approach for the timely identification of skin cancer through the classification of dataset photos into benign or malignant categories. The dataset comprises of 2437 training photographs, 660 samples, and 200 verification pictures. The classification challenge is carried out using the deep learning architectures ResNet-101 and Inception-v3. As the outcomes are scrutinized, the ResNet-101 architecture achieves an accuracy rate of 84.09%, whereas the Inception-v3 architecture achieves an accuracy rate of 87.42%. Research published in the Journal of Medical Internet Research in 2018 employed a ResNet-50 model to precisely classify skin lesions as either malignant or benign, with an accuracy of 89.9%.

Overall, deep learning models were shown to be the most accurate at identifying cancer. However, The efficacy of these models in identifying alternative manifestations of skin cancer, among which are carcinoma of basal cells and squamous cell carcinoma, was lower.

The researchers stated that more study is needed to enhance the efficacy of deep learning models for skin cancer diagnosis as well as to address concerns like dataset bias and generalizability [10].

3. Methodology

ResNet-50, ResNet-101, VGG16, and VGG19 are well-known and widely used deep learning architectures. They have been proven effective in a variety of computer vision tasks, making them solid choices as baseline models. Several papers already prove the network‘s performance on medical images [11] and based on the comparison results within the paper, ResNet and VGG can compete with state-of-the-art model. The VGGNet is a deeper and more stable architecture for ImageNet technology, which is why the VGG-19 was chosen. The evaluation of the VGGNet deep neural network (DNN)-based model for diabetic retinopathy (DR) involved the utilization of principal component analysis (PCA) and also singular value decomposition (SVD) as methods for feature selection. The results demonstrated that the VGGNet model beat both the AlexNet and SIFT models in terms of classification accuracy. Specifically, the VGGNet model achieved a classification accuracy of 92.21% for the FC7 layer and 97.96% for the FC8 layer, whereas the AlexNet and SIFT models achieved classification accuracies of 98.34% and 98.13% for the FC7 layer, respectively. Despite encountering training and localization challenges, the VGG network has exceptional accuracy in classifying images of significant scale. In response to this challenge, they propose an alternative methodology that has the potential to resolve the problem at hand. The researchers integrate a ResNet network into the VGG architecture in order to minimize training error and enhance the capability to detect objects of smaller sizes [12] [13].

3.1 Residual Network Architecture

In this paper, we utilize the ResNet-50 and ResNet-101 deep learning networks with 50 and 101 layers, respectively, for skin cancer classification. ResNet was invented in 2015 and won the classification challenge in the ImageNet competition. ResNet proposes residual connections, introduces previous training information, alleviates gradient disappearance, and effectively improves the neural network layers, which leads to increased accuracy [14]. Fig.1 shows the bottleneck design of ResNet-50 and ResNet-101. ResNet-50 consists of 48 layers and a fully connected layer for classification, so 1 + 48 + 1 = 50. ResNet-101 consists of 99 convolution layers with 7x7 and 3x3 kernel sizes and the last output layer with a SoftMax activation function, so 1+ 99+ 1 = 101. The reason for using SoftMax is its capability for multiclass classification. In the training stage, we use Adam optimizer with 100 epochs and fixed-size 224 x 224 input images.

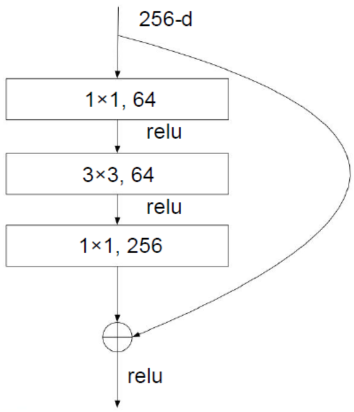


Fig. 1. The bottleneck of ResNet-50 and ResNet-101.

3.2 VGG Network Architecture

We also utilize the convolutional neural network that has 16 and 19 layers, which is called VGG16 and VGG19 Net, in pancreatic cystic type classification. VGG19 was invented by the Visual Geometry Group (at Oxford University) [15].Figure 2 shows the VGG network architecture used in our work. The VGG16 architecture is comprised of a total of thirteen convolutional layers, each utilizing a kernel size of 3 x 3. Additionally, there are five max pool layers, each employing a pool size of 2 x 2. The final layer of the network is an output layer that utilizes the softmax activation function. The VGG19 architecture consists of 16 layers of convolution, with each layer employing a kernel size of 3 × 3. Additionally, there are five maxpool layers with a pool size of 2 x 2. The last layer of the network employs a softmax activation function. The rationale behind employing the softmax function is its ability to effectively handle multiclass classification tasks. During the training phase, the Adam optimizer is employed, utilizing 100 epochs and fixed-size input images of dimensions 224 x 224.

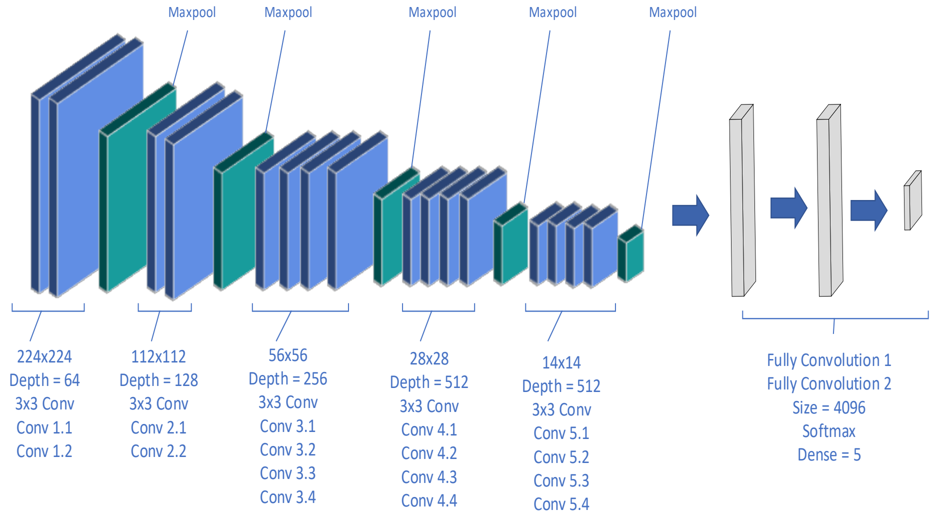


Fig. 2 Architecture of VGG Network

3.3 Dataset

We utilized a comprehensive dataset of 3,297 dermatoscopic skin cancer images, comprising 1760 benign lesions and 1457 malignant lesions and 80 test images. We obtained the skin cancer images from the public dataset proposed by [16]. This dataset has two classes: benign and malignant. In this research, we used the Windows 10 operating system with an CPU Intel Core i5-6500 @3.2 GHz and 8 GB of DDR4 memory. We also use Visual Studio 2019 as the software platform, with OpenCV library and Python programming language. After the training, all models were evaluated using 80 test images.

Table 1. Dataset information

Dataset	Images
Training	2557
Validation	660
Testing	80

4. Experimental Result

During the testing phase, the trained ResNet-50, ResNet-101, VGG16, and VGG19 networks are utilized to successively process each test image. The determination of the skin cancer kind of an input image can be made based on the maximum score obtained from the two values at the output layer. The classification outcome for the type of skin cancer is indicated by the text displayed below the test image. The evaluation of the model was conducted by considering accuracy, precision, recall, and F1-score, as these metrics hold significant clinical relevance in accurately detecting malignant lesions. The experimental results of the test photographs are presented in Table 2. The presented data illustrates the comparative outcomes of the four networks with respect to accuracy, precision, and recall.

Table 2. Four networks result in test stage

Dataset	Accuracy	Precision	Recall	F1 Score
ResNet50	0.7875	0.79	0.83	0.77
ResNet101	0.75	0.76	0.80	0.74
VGG16	**0.8375**	**0.84**	**0.85**	**0.84**
VGG19	0.7375	0.74	0.75	0.74

Table 2 presents the comparative outcomes in the classification of skin cancer across four successive networks. The benchmarking outcomes for the four models are presented below. The ResNet50 model achieved correctness, precision, recall, and F1 Score scores of 0.7875, 0.79, 0.83, and 0.77, severally. The ResNet101 model has accuracy, precision, recall, and F1 Score values of 0.75, 0.76, 0.80, and 0.74, severally. In terms of accuracy, precision, recall, and F1 Score, VGG16 demonstrates the highest scores of 0.8375, 0.84, 0.85, and 0.84, severally. The VGG19 model achieves accuracy, precision, recall, and F1 Score scores of 0.7379, 0.74, 0.75, and 0.74, severally. In addition to the evaluation score, we also obtained the confusion matrix from the four networks to further corroborate our findings. The confusion matrix for each of the networks is depicted in Figure 3.

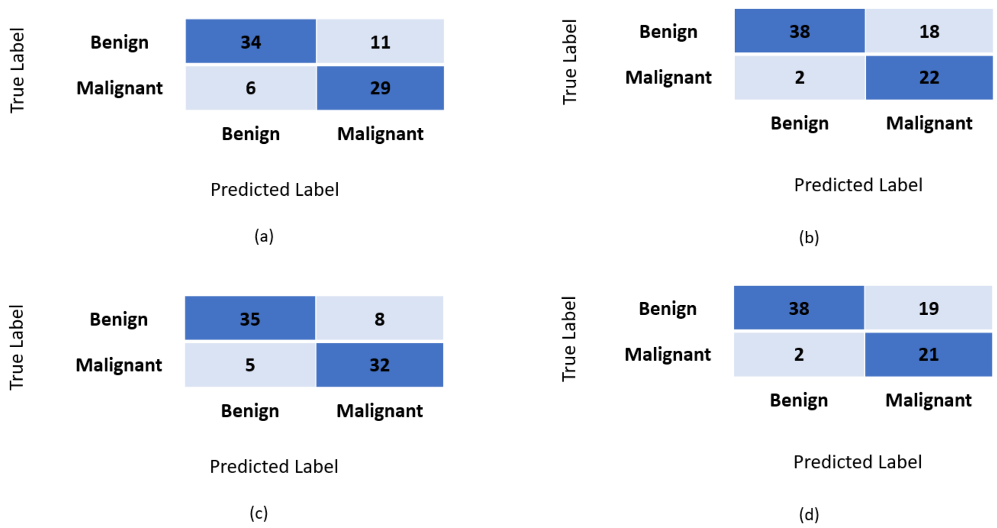
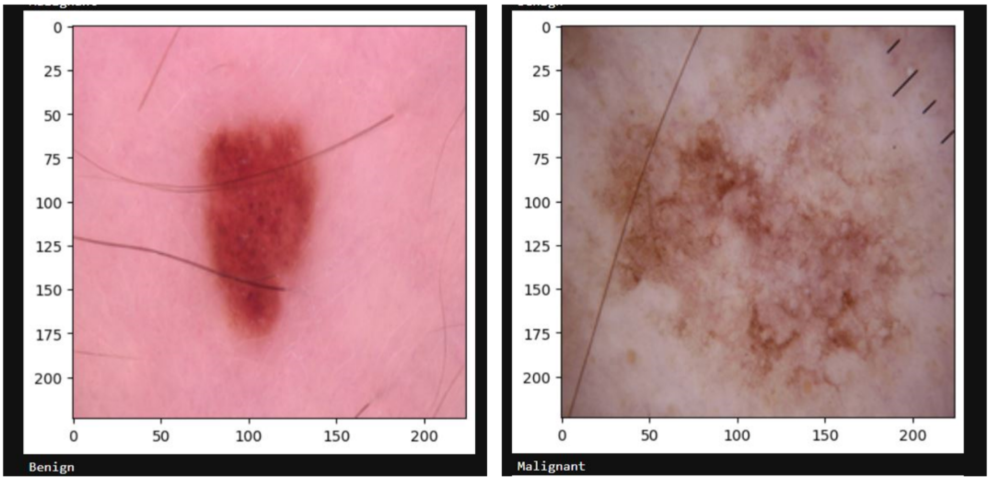


Fig. 3 Confusion matrix from (a) ResNet50 (b) ResNet101 (c) VGG16 and (d) VGG19

During the evaluation phase, the prediction result was also printed out. The class with the highest prediction value score will be used as the prediction class result, and it will be printed below the test image. Figure 4 depicts an illustrative instance of a projected examination image within the categories of benign and malignant.



(a) (b)

Fig. 4 Prediction result (a) Benign (b) Malignant

All four models demonstrated reasonably high accuracy in skin lesion classification. VGG16 achieved the highest accuracy at 83.75%, followed by ResNet50 at 78.7%. However, these differences in accuracy are not substantial, and other metrics need consideration. VGG16 and ResNet50 achieved the highest recall scores, 85% and 83%, respectively, indicating their ability to identify a higher proportion of malignant lesions. VGG16, which outperforms ResNet50, is a significantly shallower model. This raises no concern about computational resources and inference time in practical clinical applications.

5. Conclusion

This paper utilizes the deep learning neural networks ResNet-50, ResNet-101, VGG16, and VGG19 with an accuracy of 78.75%, 75%, 83.75%, and 73.75% respectively. ResNet-50 can correctly classify 63 images from 80 test images; ResNet-101 can correctly classify 60 images from 80 test images; meanwhile VGG16 can correctly classify 67 images from 80 test images; and VGG19 can correctly classify 59 images from 80 test images. Noted that we did not use any data preprocessing, but this research remains competitive and promising to look for. In summary, VGG16 emerged as a strong contender for skin cancer classification with high accuracy and balanced precision-recall trade-offs. We believe this network can be enhanced further. In the future, more images for training and testing stages are required for further improving this proposed method, and we will tailor some image preprocessing and postprocessing into it.

References

[1] A. W. Setiawan, "Effect of Color Enhancement on Early Detection of Skin Cancer using Convolutional Neural Network," in *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*, Doha, Qatar: IEEE, pp. 100–103, Feb. 2020.
[2] M. R. Hasan, M. I. Fatemi, M. Monirujjaman Khan, M. Kaur, and A. Zaguia, "Comparative Analysis of Skin Cancer (Benign vs. Malignant) Detection Using Convolutional Neural Networks," *J. Healthc. Eng.*, vol. 2021, pp. 1–17, Dec. 2021.
[3] B. P. Vrigazova, "Detection of Malignant and Benign Breast Cancer Using the ANOVA-BOOTSTRAP-SVM," *J. Data Inf. Sci.*, vol. 5, no. 2, pp. 62–75, Apr. 2020.
[4] N. Aburaed, A. Panthakkan, M. Al-Saad, S. A. Amin, and W. Mansoor, "Deep Convolutional Neural Network (DCNN) for Skin Cancer Classification," in *2020 27th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, Glasgow, UK: IEEE, pp. 1–4, Nov. 2020.
[5] K. Sambyal, S. Gupta, and V. Gupta, "Skin Cancer Detection Using Resnet," *SSRN Electron. J.*, 2023.
[6] S. Kalouche, "Vision-Based Classification of Skin Cancer using Deep Learning".
[7] Md. M. R. Dipu Chandra Malo and M. M. K. Jahin Mahbub, "Skin Cancer Detection using Convolutional Neural Network," in *Computing and Communication Workshop and Conference (CCWC)*, pp. 0169–0176, Mar. 2022.
[8] N. Abuared, A. Panthakkan, M. Al-Saad, S. A. Amin, and W. Mansoor, "Skin Cancer Classification Model Based on VGG 19 and Transfer Learning," in *2020 3rd International Conference on Signal Processing and Information Security (ICSPIS)*, DUBAI, United Arab Emirates: IEEE, pp. 1–4, Nov. 2020
[9] A. Demir, F. Yilmaz, and O. Kose, "Early detection of skin cancer using deep learning architectures: resnet-101 and inception-v3," in *2019 Medical Technologies Congress (TIPTEKNO)*, Izmir, Turkey: IEEE, pp. 1–4, Oct. 2019.
[10] N. C. F. Codella *et al.*, "Skin lesion analysis toward melanoma detection: A challenge at the 2017 International symposium on biomedical imaging (ISBI), hosted by the international skin imaging collaboration (ISIC)," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, Washington, DC: IEEE, pp. 168–172, Apr. 2018.
[11] A. Victor Ikechukwu, S. Murali, R. Deepu, and R. C. Shivamurthy, "ResNet-50 vs VGG-19 vs training from scratch: A comparative analysis of the segmentation and classification of Pneumonia from chest X-ray images," *Glob. Transit. Proc.*, vol. 2, no. 2, pp. 375–381, Nov. 2021.
[12] M. Mateen, J. Wen, Nasrullah, S. Song, and Z. Huang, "Fundus Image Classification Using VGG-19 Architecture with PCA and SVD," *Symmetry*, vol. 11, no. 1, p. 1, Dec. 2018.
[13] M. F. Haque, H.-Y. Lim, and D.-S. Kang, "Object Detection Based on VGG with ResNet Network," in *2019 International Conference on Electronics, Information, and Communication (ICEIC)*, Auckland, New Zealand: IEEE, pp. 1–3, Jan. 2019.
[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA: IEEE, pp. 770–778, Jun. 2016.
[15] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition." arXiv, Apr. 10, 2015. Accessed: Sep. 11, 2023.
[16] D. Gutman *et al.*, "Skin Lesion Analysis toward Melanoma Detection: A Challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC)." arXiv, May 04, 2016. Accessed: Jul. 28, 2023.