SKIN CANCER CLASSIFICATION

Zhe Cheng

School of Electrical Engineering, Computing and Mathematical Sciences

Curtin University

This report is submitted for COMP6016 Computer Science Project 2

June 2024

DECLARATION

This document is the result of my own work and includes nothing, which is the outcome of work done in collaboration except where specifically indicated in the text. It has not been previously submitted, in part or whole, to any university of institution for any degree, diploma, or other qualification.

Signed:
ZheCheng____

Date:12/06/2024____
____

Zhe Cheng (Master of Computing – Artificial Intelligence Major)

Curtin University

ABSTRACT

Skin disease is a big global challenge that influences all populations. As the body's biggest organ, the skin can play an important role in resisting environmental aggressors and actively participate in various physiological processes. However, the skin can easily get diseases, from inflammatory diseases to infectious diseases and serious malignancies such as skin cancer. The complexity of skin disease diagnosis comes from the complex features of the skin and subtle manifestations of early-stage disease. Early manifestations of skin diseases can easily go unnoticed, leading to delayed treatment and worse outcomes.

This research explores integrating deep learning technologies in skin disease diagnosis to improve accuracy and efficiency, which can enhance traditional skin disease diagnosis. Traditional methods of diagnosing skin disease rely on dermatologist expertise. Specifically, the method is to adjust the ResNet model better to fit the unique features of skin cancer images and provide reliable diagnosis tools to medical professionals. My research contains several tasks, including customizing ResNet's architecture to better capture skin cancer features, comparing the performance of a linear classifier and the adjusted ResNet model, and enhancing feature utilization through advanced layers in ResNet 18. This research aims to prove that a tailor-made ResNet model can reduce the subjectivity and variability in skin disease diagnosis and leverage computational power to achieve high accuracy and efficiency in medical image analysis. This paper outlines the approach, methodologies, and anticipated contributions to dermatological diagnostics.

CONTENTS

1 INTRODUCTION

1.1 Background Information

Skin disease is an important global health problem, affecting people of different ages, races and geography [1]. As the human body's largest organ, the skin protects against environmental factors and plays an important role during various physiological processes [1]. However, the skin is easy to suffer from attacks from different diseases, including inflammatory conditions, infectious diseases, and malignancies such as skin cancer [1]. For medical professionals, detecting skin disease is a challenging task.

The complexity of skin and the diverse forms of skin cancer expression need expertise and experience [2]. Various factors, such as skin texture, pigmentation, and hair colour, can increase the complexity of the diagnostic process [2]. Furthermore, early symptoms of some skin diseases look insignificant; they are easily ignored, which may lead to delayed diagnosis and worse disease conditions [3]. Accurate and timely diagnosis is important to sufficient treatment and manage skin cancer disease, especially in malignant conditions such as skin cancers [2]. Misdiagnosis or delayed treatment may lead to irreversible consequences, highlighting the importance of improving the efficiency and accuracy of developing advanced skin disease diagnostic methods [2].

To reduce the misdiagnosis from manual examination and increase the efficiency of diagnosis, researchers started to identify skin cancer using human-computer collaboration technologies. Automatic computer classification technology can assist dermatologists in making a diagnosis while keeping the objective, which reduces the workload of dermatologists and improves the accuracy of the diagnosis [4]. However, many things could be improved with computer classification technology. For example, there are high similarities between skin lesions by comparing figure 1.1 and 1.2 in colours, shapes and textures. Even between two of the same categorical skin lesions, variability still exists. Next, in the skin images, the size of the lesion areas is different. Part of the lesions take up a small part of the picture; its line of demarcation from normal skin is blurred. Thirdly, excess distracting information exists in the skin image, such as artifacts like hair, artificial labelling, and texture, making it difficult for the computer to classify skin cancers accurately.

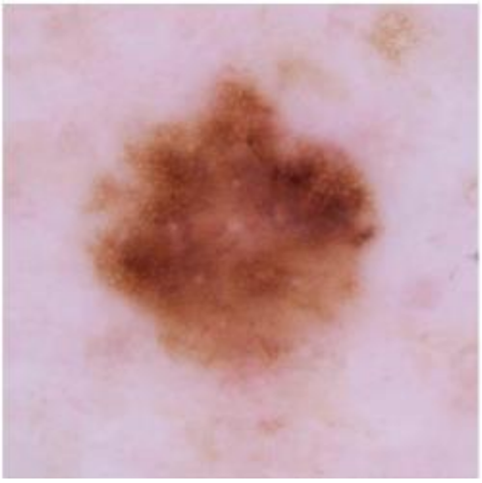


Figure 1.1 malignant melanoma

Figure 1.2 benign nevus

1.2 Research Task

Overall Objective

This research aims to use deep learning technology to improve the accuracy and efficiency of skin cancer classification. By leveraging the ability of modern computational models, this research seeks to overcome the limitations of traditional diagnostic methods, providing stable tools for dermatologists.

Task One: ResNet Model Adaptation for skin cancer image classification

Objective: Custmise the architecture of ResNet better to adapt the specific features of skin cancer images, improving its effectiveness in the medical image analysis

Methodology: Adjust the image dataset, including diverse examples of skin cancer. Adjust the architecture of ResNet to fit changes, especially optimising the output layer to reflect different categories in the skin cancer dataset.

Expected Results: Compared with the standard model, a tailor-made ResNet model can accurately classify the skin cancer images.

Task Two: Comparative Analysis of Classification Methods

Objective: Evaluate the performance of a simple linear classifier and adapted ResNet model on the skin cancer dataset and analyse their advantages and limitations.

Methodology: Use the same dataset to train a linear classifier and the adapted ResNet model and compare their accuracy and performance differences, identifying each model's advantages.

Expected Results: Compare the advantages and limitations of deep learning models and traditional machine learning methods in handling complex image data.

Task Three: Enhanced Feature Utilisation with ResNet 18 for Skin Cancer Classification

Objective: Leveraging ResNet's advanced feature-extracting ability to improve classification performance further.

Methodology: Extract features from the last convolutional layer of the ResNet 18 model, use extracted features to train a softmax classifier, and then check whether the classification performance has been improved.

Expected Results: Integrating features from ResNet 18 is expected to improve the classification accuracy.

2 REVIEW OF LITERATURE

2.1 Background

As globalisation and environmental changes have an increasing impact on human health, skin disease has become a common health problem affecting millions of people worldwide. Skin disease has a significant impact on the individual's life and brings big challenges to the public health system. Due to the wide variety of skin diseases and their varying symptoms, traditional diagnostic methods based on vision and experience usually exist at the risk of misdiagnosis and missed diagnosis. Therefore, developing more accurate and automatic diagnostic tools has important application prospects.

Against this background, leveraging modern image processing technology for automatic skin disease classification is important. Wei, Gan, and Ji (2018) proposed a method that combines gray-level co-occurrence matrix (GLCM) and support vector machine (SVM) classification techniques, efficiently improving the diagnostic accuracy of specific skin diseases such as herpes, dermatitis and psoriasis [5]. This method not only accurately extracts features of colour and texture from skin images but also improves objectivity and accuracy of diagnosis by machine learning algorithms for effective lesion classification.

Otherwise, followed by advances in artificial intelligence and deep learning technologies, image analysis methods are expected to be further optimised, including more complex algorithms and comprehensive data sets, to identify more types of skin diseases. It is not only helpful in improving the speed and accuracy of diagnosis, but it can also realise early detection and treatment, finally improving patient treatment outcomes and quality of life.

Through this method, classifying and diagnosing skin diseases using image processing and machine learning techniques is not only the performance of technological innovation but also a necessary step in medical practice to deal with the challenges of complex diseases. Therefore, continuing in-depth research and technology development in this area will directly influence the direction and effect of future skin disease diagnosis and treatment.

2.2 Traditional skin disease detection methods

Digital image processing technologies are important in traditional skin disease detection methods. Oyola (2012) found an important early detection tool that combines different image processing techniques, such as colour transformation, histogram equalization, edge detection, and circular Hough transforms, efficiently identifying chickenpox rashes and using statistical tests to reduce false positives. [6] On the other hand, Chung and Guillermo (2000) developed a system based on partial differential equations (PDE) for enhancing the image contrast, removing the hair interference in images and precisely segmenting skin lesions by geodesic active contours or edge tracing techniques that show the potential of advanced image processing technology in the skin disease detection. [7] Although these traditional methods are efficient, they rely on complex preprocessing and user participation, highlighting the importance of developing more automated, user-friendly, and precise technologies for wider application in clinical settings.

2.3 Traditional machine learning methods

Researchers have developed multiple effective techniques and models in traditional machine-learning methods for skin disease detection. Sumithra et al. (2015) introduced an automatic skin disease segmentation and classification method. This method can remove unneeded hair and noise by filtering first and then executing a regional growth approach to initialise seed points for segmentation. Afterwards, the extracted colour and texture features of skin lesion areas are classified using support vector machine (SVM) and K-nearest neighbours (k-NN) classifiers and their fusion. The designed system was tested on 726 samples, including five categorical diseases, showing good performance. [8]

Zara Naeem et al. (2022) describe a healthcare model for predicting skin cancer using deep extreme machine learning. This healthcare model utilises SVM to categorise image processing strategies, using multiple pre-processing strategies to remove noise and enhance images, aiming to improve the efficiency of early detection by thresholding and GLCM methods. [9]

2.4 Deep learning methods

In recent years, the progress of deep learning techniques utilised the power of diffusion-convolutional neural networks (DCNN) and transfer learning to improve diagnostic accuracy. Hosny, Kassem and Foaud（2018）used pre-trained DCNN and transfer learning to develop an automatic skin cancer classification method. The classification accuracies of three categorical lesions, including melanoma, common nevus, and atypical nevus, have been improved by applying fine-tuning and data augmentation to the AlexNet and replacing the last layer with the softmax layer. Their accuracies achieved 98.61 per cent, 98.33 per cent, 98.93 per cent and 97.73 per cent, respectively. [10]

Otherwise, Akter et al. (2022) proposed a multi-class skin cancer classification architecture to leverage multiple transfer learning models, including Resnet-50, VGG-16, Densenet, Mobilenet, Inceptionv3 and Xception. They applied these models to the public HAM10000 dataset, classified seven skin diseases, and compared and analysed the results. Their method proved the effectiveness of deep learning models in identifying cancer and non-cancer cells. [11]

These researches enhanced the revolutionary impact of deep learning technology in dermatology diagnosis, providing tools that improve accuracy and reduce diagnostic errors, thereby having the potential to save lives through earlier and more accurate detection.

3 METHODOLOGY

3.1 Methodology

ResNet

Residual Network (ResNet) is deep learning that effectively handles deep neural networks. ResNet introduces a new architecture called ‘residual blocks’ where layers learn the difference between input and output rather than learning the output directly. [12]

Convolutional neural network (CNN)

CNN can effectively extract and classify the features of images by its structure designed specifically for image data, including convolutional layers, pooling layers, and fully connected layers. [13]

3.2 Approach

Task One

After initiating the ResNet 18 model using ‘torch-vision.models”, I changed the output layer of the ResNet model to match seven skin cancer categories in the HAM10000 dataset. With a ratio of 8:2, I organised the image dataset into a training set and test set to validate the model’s performance on unseen data. The modified ResNet 18 model used the cross-entropy loss function for multi-class classification and the Adam optimiser to learn the weights.

Task Two

I define a custom model class 'LinearSoftmaxModel' with a single linear layer. The design of this single linear layer is to directly map the input flattened image vector to class probabilities without defining an activation function within the model. ‘CrossEntropyLoss’ has been selected as a loss function, It contains softmax operation internally. ‘CrossEntropyLoss’ calculates the cross-entropy loss between the logits output by the model and the true labels while converting the logits into probabilities.

3.3 Data Collection

I used the ‘Human Against Machine with 10,000 training images (HAM10000)’ dataset as my skin cancer classification training data. This dataset is a dermoscopic image set designed to improve the training and accuracy of automated diagnosis systems from two different sources ‘Australia’ and ‘Asuria’, spanning two decades. At the Australia site, image data was initially collected in analog and digital formats as technology advanced. At the Asuria site, image data was collected using a digital system in PowerPoint and Excel format at the beginning. [14]

3.4 Treatment of Data

In the data preprocessing step, I used the Histogram Equalization (HE) method to enhance the contrast of skin cancer images. The HE method is a simple and efficient image contrast enhancement method, which can change the image's contrast by adjusting the image's grey level distribution, making gray levels more evenly distributed throughout the image [15]. Using the HE method, the details of the skin cancer image can be preprocessed more clearly, especially when the boundaries between dark and light areas are more obvious.

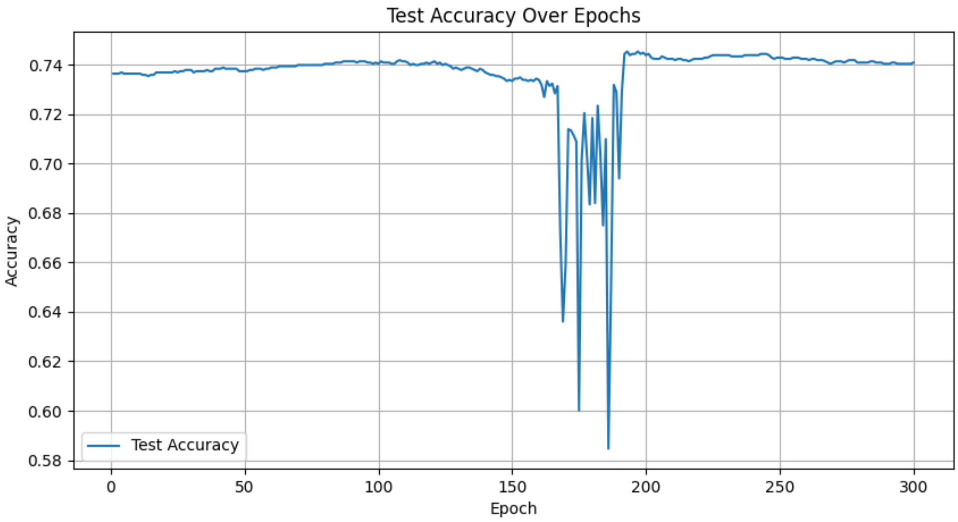
4 Results

4.1 Task One

Histogram Equalization preprocessing has little impact on model training time. However, the model's accuracy was reduced after using HAM10000 preprocessed by the Histogram Equalization method, which is strange and needs to be researched in future work. From the training accuracy curve, the model's accuracy tends to be stable after 250 epochs, and the best accuracy reaches 75 percent.

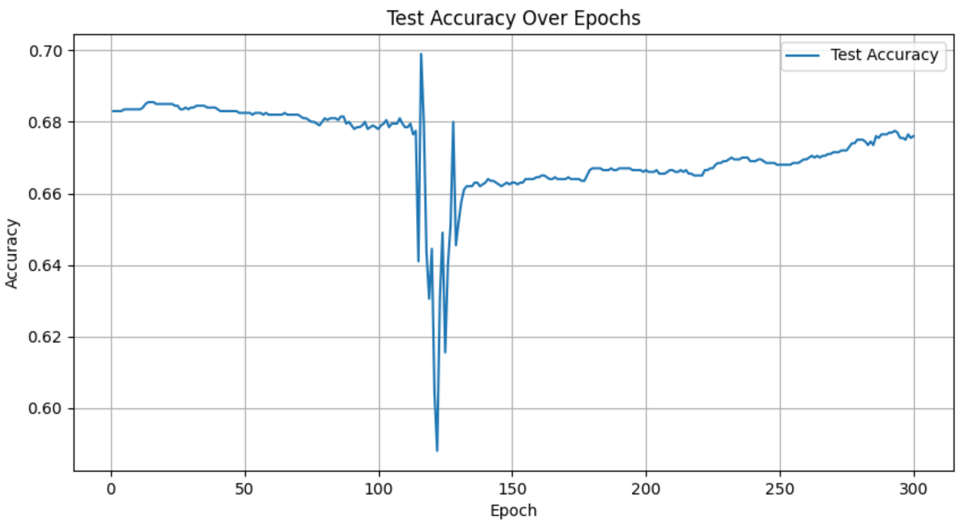
4.1.1 Unprocessed dataset test results

epoch	**test accuracy**	**training time cost (s)**
10	0.68597	271.86512565612793
50	0.59311	1363.8210141658783
100	0.64653	2706.9831895828247
120	0.75387	3255.759022474289
140	0.73789	3792.2910594940186
160	0.75287	4339.445217370987
200	0.75886	5424.420564889908
220	0.73590	6027.460168838501
250	0.74788	6851.412352561951
280	0.74588	7619.529524803162
300	0.74089	8174.329733610153



4.1.2 Test results on preprocessed dataset

epoch	test accuracy	training time cost
10	0.68	271.86512565612793
50	0.64953	1373.4158470630646
100	0.65302	2740.6628704071045
120	0.64603	3256.2995297908783
140	0.64503	3807.753484249115
160	0.65452	4352.047519207001
200	0.68298	5466.173119068146
220	0.66800	6024.548757791519
250	0.66350	6870.979373216629
280	0.65252	7601.568706274033
300	0.67599	8194.592379570007

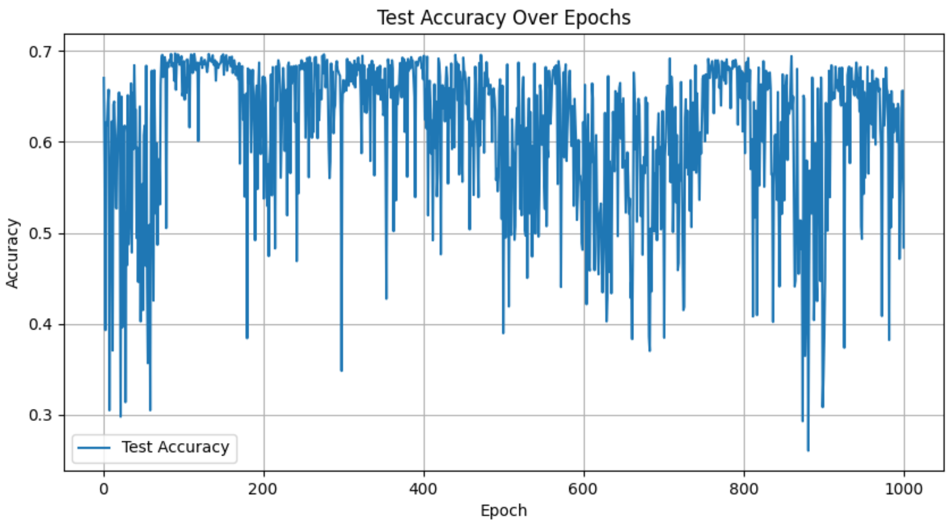


4.2 Task Two

From the perspective of training time, the preprocessing method of Histogram Equalization does not effectively reduce the training time of the model. From the perspective of training time, the linear model has a shorter training time, but the accuracy is very unstable and does not perform well on the HAM10000 skin disease data set. The classification accuracy of the Resnet model is better than the linear classifier model.

4.2.1 Unprocessed dataset test results

epoch	ResNet test accuracy	Linear Regression test accuracy	ResNet test training time	Linear Regression training time (s)
10	0.68597	0.30155	271.86512565612793	106.01637935638428
50	0.59311	0.67898	1363.8210141658783	528.388298034668
100	0.64653	0.35247	2706.9831895828247	1066.3126821517944
120	0.75387	0.63255	3255.759022474289	1297.180364370346
140	0.73789	0.55317	3792.2910594940186	1498.0130066871643
160	0.75287	0.17524	4339.445217370987	1718.6990826129913
200	0.75886	0.67898	5424.420564889908	2150.5006716251373
220	0.66800	0.53570	6024.548757791519	2390.9182074069977
250	0.66350	0.67848	6870.979373216629	2699.670417070389
280	0.65252	0.65052	7619.529524803162	3075.172013759613
300	0.74089	0.52172	8174.329733610153	3294.4216129779816
600	_	0.54219	_	6560.00859951973
800	_	0.54718	_	8926.219183206558
1000	_	0.48377	_	11070.43605542183



4.2.2 Test results on preprocessed datasets

epoch	**ResNet test accuracy**	**Linear Regression test accuracy**	**ResNet test training time**	**Linear Regression training time (s)**
10	0.68	0.65102	271.86512565612793	105.97482776641846
50	0.64953	0.60559	1373.4158470630646	528.9391100406647
100	0.65302	0.67798	2740.6628704071045	1063.1777305603027
120	0.64603	0.39790	3256.2995297908783	1285.1929399967194
140	0.64503	0.66400	3807.753484249115	1492.9226875305176
160	0.65452	0.66001	4352.047519207001	1715.2991197109222
200	0.68298	0.67099	5466.173119068146	2158.2739305496216
220	0.66800	0.65402	6024.548757791519	2347.6743919849396
250	0.66350	0.61158	6870.979373216629	2741.075968503952
280	0.65252	0.57664	7601.568706274033	3029.1770174503326
300	0.67599	0.60110	8194.592379570007	3233.675632953644
600	_	0.66001	_	6620.478531360626
800	_	0.64753	_	8697.475984334946
1000	_	0.64054	_	11095.01648569107

5 CONCLUSION

This research examined the deep learning models, especially the performance of the adjusted ResNet model compared to linear classifiers in classifying skin cancer images. This research enhanced the potential of deep learning to improve the diagnostic process of dermatology. The performance of the adjusted ResNet model is better than the performance of the linear classifier, which represents the advanced capability in handling complex image data.

Interestingly, applying histogram equalisation preprocessing seems to reduce the model's accuracy. Although histogram equalisation usually aims to enhance the contrast of the images, it may interfere with the model's ability to recognise subtle differences in images of skin lesions. This phenomenon is obvious in the processed dataset, where model accuracy decreased compared to the unprocessed dataset. Further investigation is needed into optimal preprocessing techniques for skin cancer image datasets.

Furthermore, the training time was not reduced after applying histogram equalisation preprocessing, which represents that the histogram equalisation method did not effectively optimise the training efficiency of the ResNet model and linear classifier. However, although the accuracy is low, the short training time of the linear classifier shows its potential in situations with limited computing resources.

To sum up, the deep learning model, especially the ResNet model, has the potential for accurate skin disease classification. However, careful consideration is required in choosing preprocessing methods and the trade-offs between model complexity and training efficiency. Future research should explore alternative preprocessing techniques and refine the deep learning architecture for maximising the diagnosis accuracy while minimising the resources consumed. The continuous updates of the computational model are expected to improve the accuracy and efficiency of skin cancer diagnosis, leading to better patient outcomes.

6 References

[1] Karimkhani, Chante, Robert P. Dellavalle, Luc E. Coffeng, Carsten Flohr, Roderick J. Hay, Sinéad M. Langan, Elaine O. Nsoesie, et al. 2017. "Global Skin Disease Morbidity and Mortality." *JAMA Dermatology 153 (5): 406.* https://doi.org/10.1001/jamadermatol.2016.5538.

[2] Goyal, Manu, Thomas Knackstedt, Shaofeng Yan, and Saeed Hassanpour. 2020. "Artificial Intelligence-based Image Classification Methods for Diagnosis of Skin Cancer: Challenges and Opportunities." *Computers in Biology and Medicine 127 (December): 104065*. https://doi.org/10.1016/j.compbiomed.2020.104065.

[3] Weng, Qing Yu, Adam B. Raff, Jeffrey M. Cohen, Nicole Gunasekera, Jean-Phillip Okhovat, Priyanka Vedak, Cara Joyce, Daniela Kroshinsky, and Arash Mostaghimi. 2017. "Costs and Consequences Associated With Misdiagnosed Lower Extremity Cellulitis." *JAMA Dermatology 153 (2): 141.* https://doi.org/10.1001/jamadermatol.2016.3816.

[4] Masood, Ammara, and Adel Ali Al-Jumaily. 2013. "Computer Aided Diagnostic Support System for Skin Cancer: A Review of Techniques and Algorithms." *International Journal of Biomedical Imaging 2013 (January): 1–22.* https://doi.org/10.1155/2013/323268.

[5] Wei, Li-Sheng, Quan Gan, and Tao Ji. 2018. "Skin Disease Recognition Method Based on Image Color and Texture Features." *Computational and Mathematical Methods in Medicine 2018 (August): 1–10.* https://doi.org/10.1155/2018/8145713.

[6] Oyola, Julián, Virginia Arroyo, Ana Ruedin, and Daniel Acevedo. 2012. "Detection of Chickenpox Vesicles in Digital Images of Skin Lesions." *In Lecture Notes in Computer Science, 583–90.* https://doi.org/10.1007/978-3-642-33275-3_72.

[7] Chung, None Do Hyun, and G. Sapiro. 2000. "Segmenting Skin Lesions With Partial-differential-equations-based Image Processing Algorithms." *IEEE Transactions on Medical Imaging 19 (7): 763–67.* https://doi.org/10.1109/42.875204.

[8] Sumithra, R., Mahamad Suhil, and D.S. Guru. 2015. “Segmentation and Classification of Skin Lesions for Disease Diagnosis.” *Procedia Computer Science 45 (January): 76–85.* https://doi.org/10.1016/j.procs.2015.03.090.

[9] Zara Naeem, Ghazala Zia, Zainab Bukhari. 2022."A healthcare model to predict skin cancer using deep extreme machine learning." *Journal of the National College of Business Administration & Economics 1.2: 23-30*

[10] Hosny, Khalid M., Mohamed A. Kassem, and Mohamed M. Foaud. 2018. “Skin Cancer Classification using Deep Learning and Transfer Learning.” *2018 9th Cairo International BiomedicalEngineering Conference (CIBEC),* December. https://doi.org/10.1109/cibec.2018.8641762.

[11] Akter, Mst Shapna, Hossain Shahriar, Sweta Sneha, and Alfredo Cuzzocrea. 2022. “Multi-class Skin Cancer Classification Architecture Based on Deep Convolutional Neural Network.” *2022 IEEE International Conference on Big Data (Big Data),* December. https://doi.org/10.1109/bigdata55660.2022.10020302.

[12] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. “Deep Residual Learning for Image Recognition.” *Computer Vision and Pattern Recognition,* June. https://doi.org/10.1109/cvpr.2016.90.

[13] O’Shea, Keiron, and Ryan Nash. 2015. “An Introduction to Convolutional Neural Networks.” *arXiv (Cornell University),* January. https://doi.org/10.48550/arxiv.1511.08458.

[14] Philipp Tschandl, Cliff Rosendahl, and Harald Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Scientific Data 5 (2018): 180161*, accessed June 11, 2024, doi:10.1038/sdata.2018.161.

[15] Wan Azani Mustafa and Mohamed Mydin M. Abdul Kader, "A Review of Histogram Equalization Techniques in Image Enhancement Application," *Journal of Physics: Conference Series 1019, no. 1 (2018): 012026*, accessed June 11, 2024, https://doi.org/10.1088/1742-6596/1019/1/012026.