Journal of Autonomous Intelligence (2024) Volume 7 Issue 1
doi: 10.32629/jai.v7i1.1156

ORIGINAL RESEARCH ARTICLE

A powerful deep learning method for skin cancer detection

Ahmed Mahdi Obaid*, Aws Saad Shawkat, Nazar Salih Abdulhussein

Department of Computer Science, Al Imam Al Adham University College, Baghdad 10011, Iraq

*** Corresponding author:** Ahmed Mahdi Obaid, ahmed.altaee1977@gmail.com

ABSTRACT

The authors discussed the issues posed by our study's inability to accurately diagnose skin cancer and distinguish between various skin growths, especially without the aid of cutting-edge medical technology and highly skilled diagnosticians. To that purpose, the authors have developed a deep learning (DL) method that can recognise skin cancer from images. This study investigates the use of a convolutional neural network (CNN) and the Keras Sequential-application programming interface (API) to detect cancer in seven different types of chronic lesions. The researchers used the HAM10000 dataset, which is openly available. 10,015 skin growth images with annotations are included in this dataset. The authors used several data pre-processing techniques after reading the data but before training our model. The authors provide pre-trained data for comparison and reliability assessments. Examples of transfer learning models that are used for comparison include ResNet50, DenseNet121, and VGG11. In the area of skin growth classification for skin cancer diagnosis, this helps in the identification of improved DL application procedures. Over 97.12% of the time, our suggested model accurately predicts the sort of skin that will develop.

Keywords: AI; CNN; DL; HAM10000; skin cancer

ARTICLE INFO

Received: 17 August 2023
Accepted: 4 September 2023
Available online: 3 November 2023

1. Introduction

Each year, millions of people all around the world are affected by cancer. According to research, cancer mortality rates can be greatly lowered by early and accurate cancer identification. Sometimes requiring a great level of competence, diagnosis is frequently inaccurate or false-negative. A computerised diagnostic system based on Deep Learning (DL) should be investigated because human diagnosis is prone to error[1,2]. A subclass of Machine Learning (ML) is DL. All of these methods use input data to extract information that is then used to classify the dataset and create predictions, which are typically accompanied by the predictions' accuracy[3,4]. Through the use of layers and nodes, DL systems are excellent at forecasting outcomes, especially if supervised learning is used. Around 2 to 3 million cases of skin cancer are found each year in the world, with one in three growths falling into this category. The figures show that the prevalence of cancer is increasing day by day[5]. The prompt and precise identification of cancer and the beginning of treatment are essential for its success. Less than 5% of patients with non-melanoma cancer pass away. Early detection and treatment are equally applicable to melanoma, however, due to the more severe nature of melanoma, fatality rates will be significantly higher, according to World Health Organization[6]. While the actual number of oncologists would only rise by around 25%, the number of oncologists required is about 40%[7]. The paper claims that there is a large amount of cancer formation. Experts are needed to keep up with the rising demand; otherwise, the lack of physicians will start to have a variety of effects on the expanding percentage of users. Despite dermatologists' practical experience, conventional procedures for diagnosing skin cancer are time-consuming, subject to error, and have a poor track record[8]. A large and varied collection of images of benign and malignant skin abnormalities may be used to train a CNN-based system that can identify skin cancer and other disorders using digital image inputs[9]. Oncologists would be able to work more effectively and accurately as a result, eliminating many circumstances in which problems could arise as a result of inaccurate or misleading data. The HAM10000 dataset[10] (human against machine with 10,000 training images), which contains 10,015 images of skin conditions and growths, was used by the authors in this study. For the suggested model, the authors used the Keras Sequential API. Layers are stacked on top of one another to create models using all of this API. Some of the layers in the suggested model include Conv2D, MaxPool2D, flatten, and dense. To achieve better results and learning rates, the authors used functions like Adam optimizer and ReduceLROnPlateau.

The authors of this paper will present the studies that are relevant to the suggested strategy in the next section. The empirical investigation, which primarily consists of dataset description and pre-processing stages, is undertaken in section 3. Annealing is a section in section 4 that mostly explains how the outcomes are improved. The results from the suggested model are analysed in section 5, and the authors wrap up the paper in section 6 with a few observations and suggestions for further research.

2. Literature review

Several studies have been published such as CNN-based skin cancer classification has been suggested by Manne et al.[11]. They demonstrated a fully automated computerised approach for classifying skin lesions. Three models—ResNet-18, AlexNet, and VGG16—were pre-trained in this work to serve as feature generators. Support vector machines are then trained using these recovered properties. A CNN architecture has been proposed by Thurnhofer-Hemsi and Domínguez[12] for the detection of skin cancer. They asserted that the outcomes from the DenseNet201 network are suitable for this use. Deep neural networks (DNN) have been suggested by Rahi et al.[13] for the diagnosis of skin cancer. They used various neural network (NN) methodologies and contrasted the outcomes. CNN has been suggested by Mohamed and EI-Behaidy[14] for better skin lesion classification. They claimed to have used the HAM10000 dataset to train the state-of-the-art DenseNet-121 and MobileNet systems. Kousis et al.[15] have investigated DL approaches and a mobile app for accurate skin cancer detection. The XGBoost, an average of the best eight DL models, and an average of 15 DL models were the three designs they recommended. The literature mentioned previously is summarised in **Table 1**.

Table 1. Literature summary.

Study	Purpose	Limitations	Metrics
Manne et al.[11]	The purpose of this work is to utilise CNNs to classify skin lesions.	Clinical data of varied ages, image sizes, genders, and skin types may be used as algorithm inputs to enhance classification quality.	The achieved accuracy is 85.19%.
Thurnhofer-Hemsi and Domínguez[12]	CNN will be used to find skin cancer.	The lack of large datasets is among the most challenging parts of creating an effective automatic classification system.	The achieved accuracy is 95%.
Rahi et al.[13]	The goal is to develop an automated computerised system for diagnosing skin problems using DL techniques.	The restriction is to increase precision.	The maximum accuracy for the ResNet architecture is 90%.

Table 1. (*Continued*).

Study	Purpose	Limitations	Metrics
Mohamed and El-Behaidy[14]	This study aims to improve the classification accuracy of skin lesions.	A well-balanced dataset is required for training to increase performance.	They obtain an accuracy of 92.7% and 91.2% on assessing images that are not viewed.
Kousis et al.[15]	The goal is to increase classification average accuracy rates.	An accurate diagnosis needs a huge dataset.	It achieved an accuracy of 92.25%.

3. Empirical study

The dataset used and the various pre-processing techniques are described in this section.

3.1. Dataset description

To train an NN-based diagnostic system, a sizable collection of sorted and tagged images is needed. On the other hand, sources that provide reliable diagnoses of high-quality dermatoscopic images are either limited in size or not trustworthy. A sizable collection of annotated images were assembled into a dataset named HAM10000 by Tschandl et al.[10]. The information in this data repository comes from various sources. Over 20 years, many population groupings were investigated using a range of data collection techniques. The Institute of Dermatology, Cliff Rosendahl's skin cancer clinic in Queensland, Australia, and the Medical University of Vienna, Austria, are noteworthy sources. Multiple cleaning and preparation procedures were utilised to ensure accuracy due to the large range of data acquired so that the images could be submitted to a specifically trained NN for testing and identification. The HAM10000 dataset dermatoscopic simulacra, a benchmark for evaluating ML algorithms, contains 10,015 records overall and is accessible through the International Skin Imaging Collaboration (ISIC) repository and the Harvard Dataverse for academic ML research. The crucial diagnostic classification for pigmented lesions or growths uses images from the seven-image collection as shown in **Table 2**. The images were supported by pathology, join in-vivo confocal microscopy, and consensus among experts. The ISIC library, a freely accessible collection of dermatoscopic images, comprises over 23,665 images as of April 2020. Melanocytic lesions and nevi make up about 90% of the archive, which is the focus of the dataset. It currently sets the bar for obtaining dermatoscopic images for study due to its enormous size, organised structure, and licence rights.

Table 2. Outline of the ISIC archive and class distribution regarding April 2020.

Class label	Abbreviation	Class	No. of images
0	AKIEC	Bowen disease.	331
1	BCC	Basal cell carcinoma.	588
2	BKL	Benign keratosis-like lesions.	1677
3	DF	Dermatofibroma.	121
4	MEL	Melanoma.	2175
5	NV	Melanocytic nevi.	18,611
6	VASC	Vascular lesions.	155
	Total		**23,658**

3.1.1. Dataset comparisons

The HAM10000 database is now accessible to everyone via the Harvard Dataverse. Images and metadata were compiled under the terms of the Creative Commons Attribution-NonCommercial-4.0 International Public License and came from the Institute of Medicine at the University of Washington. This study was a joint effort between the University of Queensland and the ViDIR Group (Vienna's Technical University-Department of

Dermatology, Medical School). The sources of openly accessible dermatoscopic images are listed in **Table 3** below.

Table 3. Summary of the HAM10000 dataset's metadata.

Design type(s)	**Image format conversion objective, data integration objective, data creation objective**
Technology type(s)	Digital curation.
Measurement type(s)	Skin lesions.
Factor type(s)	Age, diagnosis, animal body part, biological sex, diagnostic procedure.
Sample characteristics(s)	The skin of body, homo sapiens.

3.2. Libraries

Section 3.2. of the study highlights the assortment of code-based libraries that played an essential role in the preprocessing of the dataset. The authors employed a variety of libraries to facilitate these tasks, and the specific libraries utilized are enumerated below:

- Python: The authors harnessed an array of Python libraries including Numpy for numerical computations, matplotlib for data visualization, seaborn for enhanced visualizations, PIL for image processing, pandas for data manipulation, DateTime for handling date and time information, os for interacting with the operating system, cv2 for computer vision tasks, itertools for efficient iteration, tqdm for progress tracking, and glob for file path manipulation.
- PyTorch: The research integrated libraries from the PyTorch framework, encompassing Torch for core functionality, torch.autograd for automatic differentiation, torch.utils.data for data loading and manipulation, and torchvision for additional tools for vision tasks.
- Scikit-learn: The authors tapped into Scikit-Learn's libraries, leveraging sklearn.metrics for assessing model performance and sklearn.model_choice for model selection.
- Keras: For image preprocessing and augmentations, the study utilized the image.preprocessing module. Additionally, callbacks were employed for dynamic model adjustments during training. The researchers also harnessed TensorFlow and Keras.models for constructing and managing the neural network models.

By employing this array of diverse libraries, the authors ensured a comprehensive toolkit for data preprocessing and model development, thereby laying the groundwork for a robust and effective approach to skin cancer classification using neural networks.

3.3. Dataset pre-processing

The training and testing sets were split 80:20 from the total dataset. The test and train sets were standardised by the authors by subtracting their mean values and then dividing them by their standard deviation values. All numbers are normalised to [0, 1] from [0, 255]. The authors used this data for the proposed model because they wanted the implementation to work even with very low-resolution coloured images, which is why the dataset contains a compiled CSV file with all of the RGB data from the images. Then, by comparing the image name to the information in the metadata, the authors assigned the images a classification between 0 to 6 in terms of the sort of cancer they depicted similar to that in **Table 2**. After this, a small fraction (20%) of the training set was utilised as a validation set for a model of this type, and the remaining amount (80%) was used to train the model. Overfitting is prevented by using a validation dataset. The authors are not changing the weights of the network using this dataset; rather, they are confirming that any improvement in accuracy over the train set is due to data that the system does not currently have or has not been exposed to, such as a validation dataset. The NN is over-fitted and training should be halted if the precision of the train set of data increases but the validity of the validation set of data declines.

Since melanocytic and nevi make up more than half of the training data as shown in **Table 2**, the authors can see that there is a significant class imbalance in the data. The authors believe that improving the data will help to solve this problem. The authors must fictitiously increase the HAM10000 dataset to prevent over-fitting problems. The volume of the current dataset can be significantly increased by the authors. Big databases frequently contain null or incomplete values. These must be met, and the authors have identified and fulfilled these goals with their resources. The authors screened for duplicates and chose unique images since duplicate images could skew the results and reveal unexpected conclusions. The authors found 4501 different images in the database. Other pre-processing procedures comprise:

- The authors modified the images' pixel value range from [0, 255] to [0, 1], which is known as rescaling.
- The images were randomly rotated by the authors between 0 and 90 degrees and 10 degrees, leaving some vacant pixels that needed to be filled. The nearest value for the fill mode argument simply fills the space with the nearest large pixel values. This is done so that our model can learn more as it views related images from various perspectives.
- Since the model isn't always given centered images, the authors have also modified the images both vertically and horizontally.
- Additionally, the authors randomly cropped up to 20% of the images.

4. Proposed model

CNN is designed to find patterns in pixelated images[16]. They are frequently employed to evaluate data with a grid-like arrangement. There are normally three layers in a CNN. The first layer in a convolution process is called the convolution layer. The second layer is the kernel, which stands for the receptive field. Following the stride, the kernel moves over the height and width of the image. The image produced by this method is depicted in 2D by the initiation map. The reaction of the kernel at each image point is shown in the activation map. The proposed model was developed by the authors using the Keras Sequential API. With this API, models are built by piling layers on top of one another as shown in **Figure 1**. Since the models need to know what shape of input they will receive, the authors must define the input shape. The first layer is Conv2D. A kernel called Conv2D has smaller dimensions than the images it processes. The layer may now travel across the images thanks to this. The integer value that Conv2D produces can be used to calculate how many output filters were created overall by convolution. The authors have employed Conv2D layers with up to 128 filters, even though the most common number of filters in Conv2D is 32. In essence, these layers influence and alter. An additional layer that frequently follows a convolutional layer is the pooling layer (MaxPool2D). The authors can downsample (pool) images using the layer. To create the downscaled image, this layer analyses the pixels in the input images and chooses the one with the highest value. Smooth and crisp features are extracted using this method. These qualities include, for example, edges, points, and so forth. Each pixel's data can be linked to the following layer thanks to the flatten layer, which converts the images into a 1D array. It combines the majority of the traits that were derived from the layers. The authors then add data and feed it through thick layers. The results of the images after being processed by the earlier convolutional layers are unknown to the authors. Beyond intricate layers, as seen in **Figure 1**, the dense layers program makes an effort to determine the connection between all of the information supplied to it. **Figure 2** depicts the suggested model's workflow.

4.1. Annealing

This section talks about how to make the model better using:

- ReduceLROnPlateau: Reduces the learning rate when a measure stops becoming better. Models usually gain by reducing training up to a factor of 2–10 when learning becomes static. This callback monitors the quantity and lowers the learning rate if an improvement is not observed after a predetermined number of iterations. The authors noted that after around the 19th or 20th epoch, this similar learning rate decreased, leading to better outcomes.

- Setting optimizer: For the suggested model, the authors used the Adam optimizer[17]. Adam is only a DNN training approach that optimises adaptive learning rates. According to Kingma and Ba[17], the technique is ideally suited for situations that are large in terms of data/parameters, are computationally efficient, and have a low memory requirement.

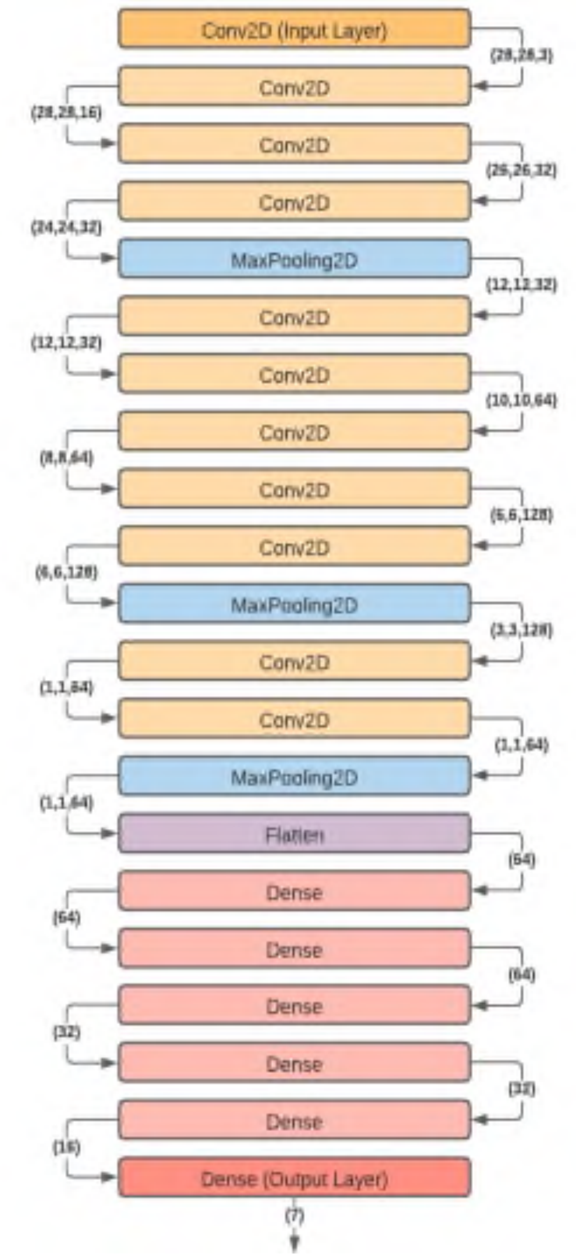

Figure 1. The design of the proposed model.

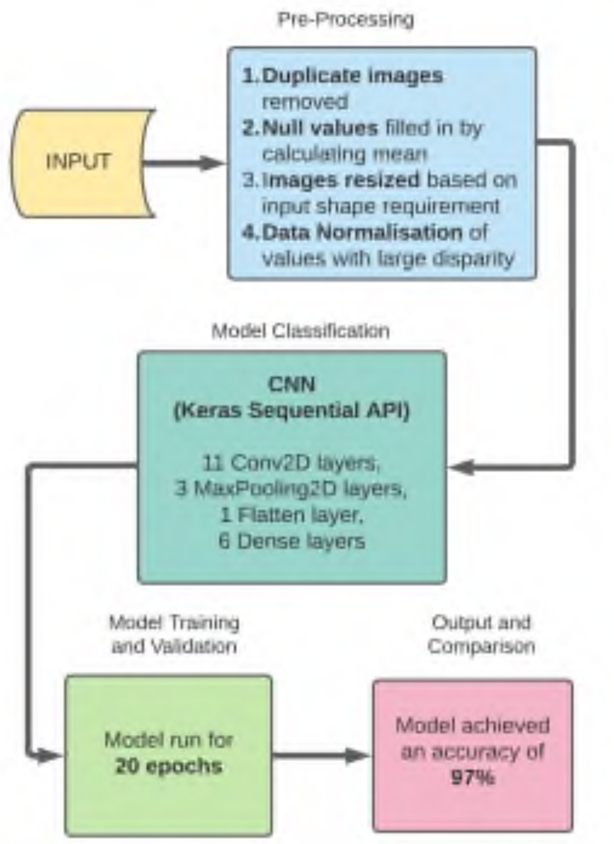

Figure 2. The proposed method's workflow.

5. Result analysis

Table 4 in the research article showcases the outcomes of the model proposed in the study, complemented by visual representations presented in **Figures 3–5**. The model underwent 20 epochs of training, ultimately achieving an impressive accuracy level of approximately 97.12%.

In line with the findings illustrated in **Table 5**, the authors conducted a comprehensive comparative analysis by employing established techniques, such as ResNet50[18], DenseNet121[19], and VGG11[20], to benchmark against the performance of previously documented models. Notably, these models are characterized by slower computation times. To facilitate a comprehensive comparison, the authors utilized the pre-trained data from ImageNet and ran it through the specified algorithms, limited to a span of 10 epochs. The accuracy results from these algorithms were then juxtaposed with the accuracy achieved by the proposed model.

According to the authors' assessment, the suggested model exhibited superior accuracy when contrasted with the three aforementioned established models. This promising outcome positions the proposed model as a potential frontrunner in practical applications, making it well-suited for real-world scenarios where industries can integrate its capabilities in conjunction with human efforts.

The substantial performance enhancement observed in the proposed model's accuracy implies its potential efficacy as a viable tool for real-world deployment. This advancement holds the promise of fostering collaboration between relevant industries and human expertise, thereby amplifying the impact of these technologies in practical contexts.

Table 4. Recall, precision, support & F1-score of the proposed model.

	Precision	Recall	F1-score	Support
AKIEE	1.00	0.88	0.92	1344
BCC	0.99	0.99	0.99	1262
BKL	0.94	1.00	1.00	1566
DF	0.98	0.99	1.00	1577
MEL	0.98	0.99	0.99	1744
NV	1.00	0.99	0.99	1211
VASE	0.99	0.99	0.99	1341
Accuracy	-	-	0.99	10,003
Macro avg.	0.99	0.99	1.00	10,002
Weighted avg.	0.99	0.99	0.99	10,002

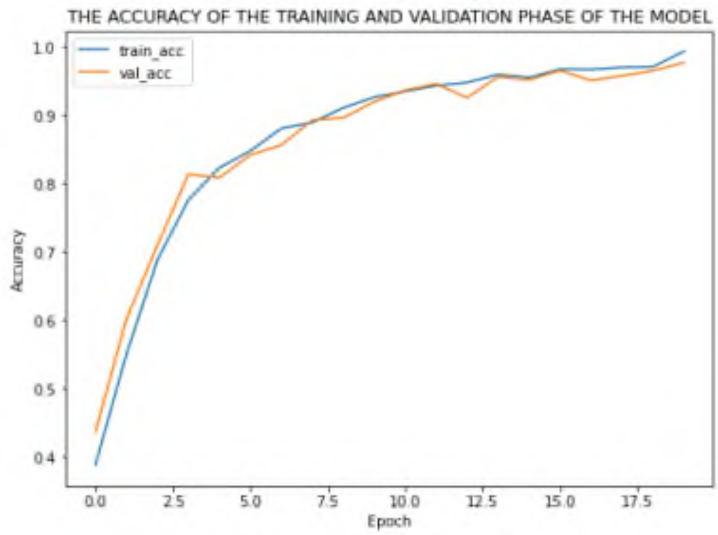

Figure 3. Accuracy of the proposed model's training and validation phase.

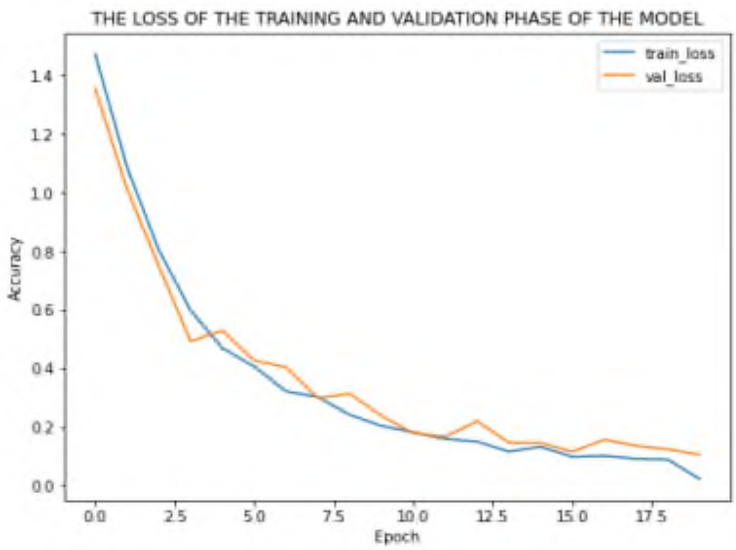

Figure 4. Loss of the proposed model's training and validation phases.

Figure 5. Using testing data, the proposed model's confusion matrix.

Table 5. Comparison of the proposed model's validation against others.

	CNN	**DenseNet121**	**ResNet50**	**VGG11**
Accuracy	96%	91%	88%	84%

6. Conclusion

Skin cancer is a highly concerning disease due to its potential to metastasize if not detected and treated promptly. In the context of this research, the aim was to leverage convolutional neural networks (CNNs) for the identification of skin lesions, a critical step in early diagnosis. The primary objectives encompassed both showcasing the feasibility of recognizing diverse forms of skin cancer from images and constructing a diagnostic system founded on a CNN model developed through the Keras Sequential API.

The CNN model demonstrated remarkable efficacy, boasting an impressive accuracy rate of at least 97.12%. This achievement positions the CNN-based skin cancer identification approach as a potentially superior tool to human visual assessment in identifying this disease. The central proposition of the study revolved around the classification of various skin cancer forms from image data using a CNN model designed with the Keras Sequential API.

The model's architecture was established with a fixed kernel size and a consistent filter size, employing a specified convolutional operation. The outcomes of this process were then condensed into 1D singular vectors, which were adopted as input variables for the Keras CNN. To enhance the dataset quality, duplicates were eliminated from the initially comprised 10,015 images.

The study's forthcoming endeavors will delve into testing deep networks and exploring alternative hierarchies, including preprocessing techniques. A rigorous examination of the attributes specific to each class will be crucial for developing a robust classifier. Furthermore, the exploration of probabilistic techniques to yield precise predictions across diverse classifiers stands as a promising avenue for further exploration.

Finally, the paper highlights the challenges involved in comparing different techniques and underscores the forthcoming issues that necessitate resolution for future progress. This research underscores the potential of CNNs in the early detection of skin cancer and underscores the multifaceted pathways for refinement and innovation in this crucial area of medical diagnosis.

Author contributions

Conceptualization, AMO and ASS; methodology, AMO; software, AMO; validation, AMO, ASS and NSA; formal analysis, AMO; investigation, AMO; resources, AMO; data curation, AMO; writing—original draft preparation, AMO; writing—review and editing, AMO; visualization, AMO; supervision, AMO; project administration, AMO; funding acquisition, ASS. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

The authors declare no conflict of interest.

References

1. Hoshyar AN, Al-Jumaily A, Hoshyar AN. The beneficial techniques in preprocessing step of skin cancer detection system comparing. *Procedia Computer Science* 2014; 42: 25–31. doi: 10.1016/j.procs.2014.11.029
2. Lau HT, Al-Jumaily A. Automatically early detection of skin cancer: Study based on nueral netwok classification. In: Proceedings of the 2009 International Conference of Soft Computing and Pattern Recognition; 4–7 December 2009; Malacca, Malaysia. pp. 375–380.
3. Seeja RD, Suresh A. Deep learning based skin lesion segmentation and classification of melanoma using support vector machine (SVM). *Asian Pacific Journal of Cancer Prevention* 2019; 20(5): 1555–1561. doi: 10.31557/APJCP.2019.20.5.1555
4. Blot M, Cord M, Thome N. Max-min convolutional neural networks for image classification. In: Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP); 25–28 September 2016; Phoenix, AZ, USA. pp. 3678–3682.
5. Skin cancer. Available online: https://www.mayoclinic.org/diseases-conditions/skin-cancer/symptoms-causes/syc-20377605 (accessed on 6 January 2023).
6. World Health Organisation. Radiation: Ultraviolet (UV) radiation and skin cancer. Available online: https://www.who.int/news-room/questions-and-answers/item/radiation-ultraviolet-(uv)-radiation-and-skin-cancer (accessed on 6 January 2023).
7. A shortage of oncologists. Available online: https://www.clearityfoundation.org/a-shortage-of-oncologists/ (accessed on 7 January 2023).
8. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017; 542(7639): 115–118. doi: 10.1038/nature21056
9. Mendonca T, Ferreira PM, Marques JS, et al. PH^2—A dermoscopic image database for research and benchmarking. In: Proceedings of the 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS); 3–7 July 2013; Osaka, Japan. pp. 5437–5440.
10. Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific Data* 2018; 5: 2018. doi: 10.1038/sdata.2018.161
11. Manne R, Kantheti S, Kantheti S. Classification of skin cancer using deep learning, convolutional neural networks-opportunities and vulnerabilities—A systematic review. *International Journal for Modern Trends in Science and Technology* 2020; 6(11): 101–108. doi: 10.46501/IJMTST061118
12. Thurnhofer-Hemsi K, Domínguez E. A convolutional neural network framework for accurate skin cancer detection. *Neural Processing Letters* 2021; 53(5): 3073–3093. doi: 10.1007/s11063-020-10364-y
13. Rahi MMI, Khan FT, Mahtab MT, et al. Detection of skin cancer using deep neural networks. In: Proceedings of the 2019 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE); 9–11 December 2019; Melbourne, VIC, Australia. pp. 1–7.
14. Mohamed EH, El-Behaidy WH. Enhanced skin lesions classification using deep convolutional networks. In: Proceedings of the 2019 IEEE 9th International Conference on Intelligent Computing and Information Systems (ICICIS); 8–10 December 2019; Cairo, Egypt. pp. 180–188.
15. Kousis I, Perikos I, Hatzilygeroudis I, Virvou M. Deep learning methods for accurate skin cancer recognition and mobile application. *Electronics* 2022; 11(9): 1294. doi: 10.3390/electronics11091294
16. Suárez-Paniagua V, Segura-Bedmar I. Evaluation of pooling operations in convolutional architectures for drug-drug interaction extraction. *BMC Bioinformatics* 2018; 19: 209. doi: 10.1186/s12859-018-2195-1
17. Kingma DP, Ba JL. Adam: A method for stochastic optimization. In: Proceedings of the ICLR 2015 International Conference on Learning Representations (ICLR); 7–9 May 2015; San Diego, CA, USA.
18. Indriyani T, Nurmuslimah S, Taufiqurrahman A, et al. Steganography on Color Images Using Least Significant Bit (LSB) Method. In: Proceedings of the 1st International Conference on Neural Networks and Machine Learning 2022 (ICONNSMAL 2022); 29–30 November 2022; Jember, Indonesia.
19. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: Proceedings

of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 21–26 July 2017; Honolulu, HI, USA. pp. 2261–2269.
20. VGG-11 architecture. Available online: https://iq.opengenus.org/vgg-11/ (accessed 7 January 2023).