Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 6, 2025

MULTICLASS SKIN CANCER CLASSIFICATION USING RESNET AND DERMOSCOPIC IMAGING

[1]Umama Ilyas,[2]Muhammad Fuzail,[3]Muhammad Kamran Abid, [4]Talha Farooq Khan, [5]Ahmad Naeem,[6]Naeem Aslam

1,2,5,6Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan

3Department of Computer Science, Emerson University Multan, Pakistan

4Department of Computer Science, University of Southern Punjab Multan, Pakistan

*Corresponding Author: Muhammad Kamran Abid (kamran.abid@eum.edu.pk)

DOI: https://doi.org/

Keywords
(Deep Learning, Skin Cancer, Skin Cancer Classification, Skin Images).

Article History
Received on 18 May 2025
Accepted on 10 June 2025
Published on 20 June 2025

Corresponding Author: *
Muhammad Kamran Abid

Abstract

Skin cancer is among the most common and life-threatening illnesses globally, and early and precise diagnosis is necessary to enhance the outcome of treatment. Conventional diagnosis is based largely on skilled dermatologists, which is time-consuming and variable. Over the past few years, deep learning methods have proven to have vast potential in computerizing medical image analysis, both speeding up and improving the accuracy of diagnosis. This work aims to develop a reliable deep learning-based classification model using the ResNet architecture for identifying and categorizing various types of skin cancer. The research makes use of the HAM10000 dataset, an extensive collection of dermoscopic images representing various skin cancer types, to ensure representative and varied data for training and testing. The methodology suggested here encompasses several stages such as dataset preprocessing, feature extraction based on pre-trained deep learning models, and classification based on ResNet. Several data augmentation methods and preprocessing operations are utilized to improve model accuracy and handle imbalances in classes. Feature extraction takes advantage of ResNet's hierarchical feature representation, where complex patterns in dermoscopic images are captured to support precise classification. The model is optimized and fine-tuned with the best hyperparameters to achieve maximum classification accuracy with minimal computational complexity. A comparative analysis with traditional machine learning methods underscores the superiority of deep learning techniques in dealing with sophisticated visual data. Experimental findings show that the model proposed has excellent classification accuracy, performing better than traditional methods in precision, recall, F1-score, and overall classification accuracy. The model also separates various types of skin cancers correctly with minimal false positives and negatives. Performance monitoring methods like cross-validation and regularization are also embedded to prevent overfitting so that the model is generalizable to unseen samples. The computational cost of the method is examined to ascertain its viability for real-time clinical use, highlighting its prospect of being part of automated diagnosis systems.

INTRODUCTION

The skin is the most important organ of the body, and it acts as the principal defence point between the outside world and the internal organs. It also exhibits a significant role in thermoregulation, sensory perception, and infection safeguard. Being regularly exposed to factors of the surrounding environment, such as ultraviolet (UV) radiation, toxins, and mechanical trauma, skin is particularly vulnerable to an enormous range of disorders, including skin cancer. Skin cancer results from to uncontrolled and abnormal multiplication of the skin cells that normally are a result of mutations of the DNA genes as caused by long exposure to the UV radiations. This with time will lead to development of malignant tumors in certain areas of the skin[1].The three most common and highly pertinent types of skin cancer include, melanoma and basal cell carcinoma (BCC) and squamous cell carcinoma (SCC). The least popular are melanoma except that this form of cancer is the most dangerous since it has the capability of spreading and claiming the lives of the sufferees. Melanoma is a type of cancer that develops in melanocytes all of which are the cells that synthesize melanin which is a pigment that makes the skin look like skin. All melanomas tend to be dark brown or black since they continue to produce melanin. BCC and SCC on the other hand are less aggressive yet more common and they originate in the basal as well as squamous cells respectively. Timely and appropriate differentiation of these skin cancers can be essential in selection of effective modes of treatment and improvement of patient prognosis[2].

Dermoscopy-guided biopsy is the classical method to diagnose skin cancer through direct visualization of dermatologists and further identification through biopsy. The human based diagnosis is prone to errors which include a subjective view and inconsistency in personnel skill and agent sickness, even though a very high accuracy can be achieved by advanced dermatologists[3]. Moreover, the efficient and timely diagnosis is very challenging especially in locations with limited access to specialist medical staff. These limitations provoke the necessity to introduce automated, reliable and scalable tools to support the process of early detection and classification of skin cancers[4]. Over the recent years, deep learning has proven to be an extremely useful tool in medical image analysis particularly dermatology. One example of a deep learning model that has experienced an unprecedented type of success is the convolutional neural networks (CNNS) in its abilities to generate complex patterns through dermoscopic images. They are able to make multiclass classification, i.e., they are quite precise at melanoma, BCC and SCC diagnosis and other harmless skin lesions[5]. The use of pretrained models such as ResNet together with the availability of large annotated dataset such as HAM10000, have also further enhanced the performance of these systems to the extent of matching or even exceeding the human level performance in terms of diagnostic skills[6].

This work aims to explore the use of deep learning-powered approach to skin cancer type multi class classification using dermoscopic images. The study focuses on optimal performance on the model because of effective preprocessing, feature extraction using pretrained CNNs, and robust testing. By training the model on real-world datasets and measuring the effectiveness of his model using the prevalent metrics that include accuracy, precision, recall, and F1-score, the research attempts to develop a practical model that can be applied in clinics. Lastly, the aim is not only automatizing the process of classification, but to provide interpretability, and confidence in the model predictions, thereby helping dermatologists with their decision-making process and improving patient care[7]. The chapter introduces the research by establishing the domain of problem, and importance of early diagnosis, as well as potential of AI in transformational potential of dermatological diagnosis. To present the full picture how deep learning can become a revelation in terms of skin cancer classification and detection the methodology, dataset features, experimental results, and comparison of results are provided in the following chapters. A few types of skin cancer according to Figure 1.

Figure 1 Skin Cancer Types

In the past decades the incidence of skin cancer has been skyrocketing across the globe because of the numerous environmental and demographical factors. The greatest among them is the increased exposure to the ultraviolet (UV) radiation due to heavy outdoor activities and destruction of the ozone layer. Meanwhile, there has been an amplified life expectancy which has increased in aged individuals who are more prone towards the development of skin related cancers[8]. According to GLOBOCAN 2022 survey, it is estimated that the incidences of new cancers worldwide were close to 20 million leading to close to 9.7 million deaths. Skin cancer has been one of the most commonly diagnosed types of cancer and their early discovery has been one of the major success factors that have helped patients recover better and reduce mortality[9].

An effective treatment of skin cancer is the premise of early diagnosis. The opportunity to diagnose the disease at its initial stage not only allows to provision medical assistance in time but also significantly increases survival rates along with the simplification of the treatment process[10]. In fact, dermatologists typically made their diagnosis during a visual and dermoscopic study of the topography of skin lesions to distinguish the benign and malignant forms, including melanoma, basal cell carcinoma (BCC), and squamous cell carcinoma (SCC). Such critical attributes such as asymmetry, irregular borders, color change and change in diameter as well as formation of lesions (as popularly confirmed by the acronym ABCDE) are important signs employed during traditional diagnosis[11]. Manual techniques rely greatly on the clinician's expertise, though, and also involve subjectivity—potentially resulting in diagnostic variations among medical practitioners[12]. To confirm the visually diagnosed problems, biopsies are performed which despite being accurate are invasive, time consuming, and costly. Such limitations have led to researchers trying to find non-invasive and mechanized systems that can either guide or even substitute human judgment[13]. Therefore, the application of artificial intelligence (AI), in this case deep learning (DL), to medical image has recently experienced a surge. Deep learning opens up the possibility of automatically learning some complicated patterns of large amounts of data such that the models are able to predict and classify the sorts of skin cancer with high-level accuracy[14]. Deep learning models, unlike traditional image processing algorithms whose features have to be designed by hand, automatically obtain the features and this significantly enhances performance and robustness[15].

The proposed research suggests a powerful deep learning-based diagnosis system on the dermoscopic image which will categorize various skin cancers. It works on the principles of two significant actions: image preprocessing and training of the model. Preprocessing: In the dermoscopic images, the initial step of preprocessing is evenness of brightness and contrast where the noise and abnormalities that normally occur as a result of variation in the lighting and skin colors are narrowed. The resizing and normalization of the images are done according to the size requirement, as an input to the model architecture so that they can easily be inbuilt into the training process. Preprocessing plays a great role in raising the overall accuracy of the classification models. The model is trained using dermoscopic images of the publicly available and commonly used dataset ISIC 2018, which has many thousands of labelled dermoscopic skin lesion images. In multiclass classification, they are comprised of different classes including melanoma, BCC, and SCC and benign nevi which is ideal. During training, the model acquires the ability to capture special visual patterns corresponding to each type of cancer. Four deep learning models were tested and compared, ResNet50, InceptionV3, and Inception-ResNetV2. They have been chosen based on their proven convenient work in the usefulness of computer vision and their ability to code both the low-level texture and high-level semantic in medical images.

A training strategy was followed to ensure generalization and exclude the issue of overfitting. Some of the techniques used are data augmentation and learning rate schedule as well as early stopping. Data augmentation has helped since it was a way to artificially inflate the dataset with slight shifts in rotation, shearing, and flipping, therefore, allowing the model to learn invariant features. There is early stopping that helped to end training when the performance in the validation set stagnated and this is a standard way to make sure that the model is efficient and when the computation should not be so long. The use of different training setups and parameter optimization ensured that the proposed method ensured that the model was robust to reality and accurate even with divergent inputs of images.This level of the model performance was assessed quantitatively with standard classification score calculation using metrics such as F1-score, area under the receiver operating characteristic curve (AUC), precision, recall, and accuracy, to mention but a few. These statistics are crucial to the assessment of the ability of the model to differentiate between classes, especially in a multiclass case, where one no longer seeks simply to identify cancer or no cancer, but to identify different lesion types correctly. Good results on these statistics imply that by using the model, it is possible to mitigate false positives and false negatives and, thus, the model will work well in a medical diagnostic setting.

In short, deep learning in skin cancer diagnosis is an effective development towards computerized medical decision-making. This work is a clarification of what AI can accomplish to support clinical practice by a combination of image preprocessing methods, state-of-the-art neural networks, and various training approaches. Such systems may serve as the decision-support tools of the dermatologist, optimizing the uniformity of their diagnosis and enabling the possibility to conduct early action on the patients who have no direct and immediate access to specialist care. The results of the present study have a positive outlook for future applications in a clinical setting when an AI-based tool may supplement the traditional diagnosis methods and help to mitigate the prevalence of skin cancer worldwide.

Related Work

The literature review is at the center of this study as it presents the findings on the studies that have been conducted regarding the use of Artificial Intelligence (AI) and Deep Learning (DL) in identifying and classifying skin cancer. An AI-based diagnosis in skin cancer is becoming even more possible with skin cancer being an emerging problem in the whole world, various improvements in the medical image analysis have happened to make it possible. The chapter intertwines recent information and technology in the field to know how deep learning, specifically the Convolutional Neural Networks (CNNs) has transformed the process of detecting and identifying the skin cancer. The overview focuses on the most influential models, methods, and findings and creates a detailed picture of the state of the field today.

The bulk of the literature significantly focuses on the classification task of melanoma and non-melanoma skin cancer, and there is a large application of CNN based methods in light of their outstanding performance on computer vision oriented tasks[16]. Majority of the studies are based on transfer learning where pre-trained model ResNet, VGGNet, or Inception are trained on medical image datasets ISIC and HAM10000 to improve accuracy[17]. The other approaches found in the literature reviewed in this study include ensemble, where a lot of models are used to improve the robustness as well as hybrid, which captures both contextual and details use of models by combining CNN with others[18]. These techniques have shown potential to improve the accuracy of diagnosis, though constraints still exist and pertain to the issue of generalisation of different datasets and imaging conditions[4]. The other heavily-enforced theme in the literature is the quality and availability of the annotated datasets, a major issue in the realm of medical AI. It takes a lot of high quality of annotated images to train robust deep learning models, and getting a dataset of this quality is not always possible, especially in the case of rare classes. It causes the class imbalance issues, which are addressed by many works using data augmentation or synthetic image generation through GANs[19]. In addition, explainable AI in dermatology is becoming more demanded. The importance of model interpretability to acquire the favor of the clinicians has been repeatedly emphasized in a number of studies. Techniques such as Grad-CAM, saliency maps and attention mechanisms are some of the most commonly applied to make predictions of models interpretable by the clinician[13].

To sum up, the literature review gave a thorough account of the trends, obstacles, and opportunities in the utilization of deep learning to diagnose skin cancer. It forms a strong rationale into this thesis by pointing out the weaknesses in existing methodologies and how this research would jump the next steps to remedy such. By combining the advantages of the past studies and redefining the limitations of the preceding ones, the current piece contributes to the progress of the AI-based dermatological diagnostics with the focus placed on the multiclass classification, optimization of the models, and practical usability.The tendency to attract significant attention of doctors to artificial intelligence (AI) and deep learning (DL) in skin cancer diagnosis has been observed. This new found interest has been necessitated by an increased demand of quick, accurate and low risk methods of diagnosing the conditions. Direct inspections and biopsies work well but they are also inaccurate and take a lot of time. With the help of deep learning algorithms, it is possible to detect skin cancer with the help of large libraries of dermoscopic photo data and do it more reliably, thanks to tools created on their basis. The worth of these techniques has been tested in numerous medical researches and found that they regularly deal with the same or even better performances as real dermatologists.

Methodology

The research methodology of the given work is focused on the creation of a deep learning-based skin cancer classification using the ResNet50 model. It was aimed to produce a reproducible, scalable and interpretable system that can classify seven types of skin lesions based on the images taken using a dermoscope. This was done by selecting and preprocessing the dataset, then configuring the model, training, and testing the performance of the model. The stages were carefully crafted to provide the best learning experience along with the resolution to the common issues in medical imaging, including imbalanced classes, overfitting, and intra-class similarities. HAM10000 dataset, abbreviated to Human Against Machine, with 10,000 training images, was chosen as the main dataset to be used in this project. It is one of the largest as well as widely used publicly available datasets in the field of skin lesion categorization[6]. The data have 10,015 high-resolution dermoscopic images labeled with one of seven classes (Melanoma(MEL), Melanocytic Nevi(NV), Basal cell carcinoma(BCC), Actinic Keratoses(AKIEC), Benign Keratosis-like Lesions(BKL), Dermatofibroma(DF) and Vascular Lesions(VASC)). These courses are the blend of both malignant and benign lesions on the skin, frequent in dermatology practice. Each image is presented as a JPEG file and has a description of clinical metadata of the image (the location of the lesion, confirmation technique, or the age of the patient). These images differ by resolutions and quality as it happens in the real world of clinics. One interesting different feature of this data is the imbalance of the classes, with most of them related to Melanocytic Nevi and marginalized, some of the classes that are Dermatofibroma and Vascular Lesions classes[20].

The ResNet50 deep convolutional neural network that takes advantage of residual learning was selected as a classification model. ResNet50 has the characteristics of training a very deep network, but not falling into problem with vanishing, because of the identity shortcut connections[21]. These left over connections enable the model to propagate information through layers and therefore enable learning of features even in complex training data. This experiment involved transfer learning in which ResNet50 model was trained on ImageNet data and then fine-tuned using HAM10000. The last layer of classification was changed to suit the seven output classes which represented the lesion types. A global average pooling layer was incorporated and placed right before the output layer to decrease the number of parameters and minimize the risk of overfitting The training of the model was conducted with Tensorflow and the Keras deep learning frameworks. The categorical cross-entropy loss function was implemented with the particular choice due to the fact the task was used in the multiclass classification. Adam optimizer was also selected due to its adaptive rate learning rate which helps in acceleration of convergence[22]. The training batch size was set at 32 and the initial learning rate at 0.0001 which was determined empirically. The training was done in 50 epochs using the early stopping which would stop training in case the validation loss failed to decrease after 10 consecutive epochs. This made it better to avoid overfitting and get the best results possible. The divide of the dataset into validation and training sets was performed as 80:20 to make sure that the model was trained on unseen data.

In order to measure the performance of the trained model, a number of performance measures were obtained, one of which was accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC). These measures gave a detailed report of the accuracy of classification of lesions against the seven categories of the model[13]. The measure of accuracy was taken to denote the overall correctness of prediction, whereas precision and recall determine the capacity of the model to not make a false positive and negative, respectively. F1-score was a balanced indicator of precision and recall and AUC-ROC was used to evaluate the discriminative capability of the model, particularly in the imbalanced classes scenario. The whole flow of the given methodology is shown in Figure 2.

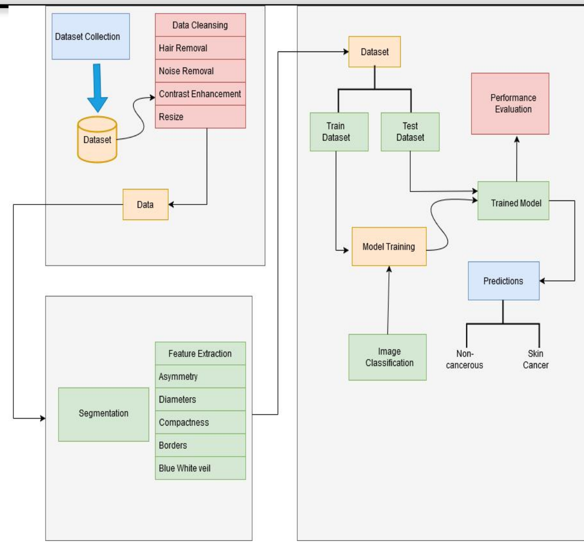

Figure 2 Flow Chart of Methodology

Besides quantitative analysis, the Grad-CAM (Gradient-weighted Class Activation Mapping) methodology was used as a method to obtain visual explanations of model outputs. Grad-CAM generates heatmaps, which indicate best best-characterized parts of input image in making the decision[13]. Such visualizations were employed in order to ensure that the model did indeed target clinically important regions to provide its classifications, e.g., the edge of lesions and pigmentation patterns. Grad-CAM can increase the model trust and facilitate its incorporation into clinical workflow by providing an interpretable output. Such visual explanations are especially pertinent in medical settings where interpreting the logic underpinning the prediction of a model is vital in the clinical decision-making process. The whole methodology was tailored in such a way that a high-performing, interpretable, and clinically useful model of skin cancer classification could be developed. All the processes such as meticulous preprocessing and augmentation, careful selection of the models and the performance assessment, led to the construction of a robust system that could help dermatologists diagnose skin lesions prematurely and correctly. The flexibility, transparency, and high-performance metrics in the model show that the model can be deployed on real-world conditions of healthcare, especially in the regions where the availability of expert dermatological attention is rare.

Results and Discussion

The assessment of the proposed deep learning ResNet50-based approach was carried out with the help of HAM10000 dataset, where dermoscopic images of seven skin lesions varieties are available. The metrics of the model that emerged after training and validation indicates the robustness of the model in the multiclass classification of skin cancer. The accuracy of the model was high at 88.25% and this implies that the model is able to learn and generalize well on different categories of skin lesion. Precision, recall, F1-score, and the Area Under the Curve (AUC) performance matrices were calculated to give a complete description of classification ability of the model. Such measures especially matter in the medical image analysis context, as false positives and false negatives may have significant clinical consequences. The model was shown to exhibit good results in all these measures, providing support to its applicability in actual diagnostic scenarios. Precision is the ratio of the number of predicted positive results that are correct, out of all of the positive predictions. It is vital in minimizing the false positive. Among the classes, the model performed very high comparisons in most of cases in identifying melanoma and melanocytic nevi. There were also high recall values indicating that the model detected a high percentage of true positive cases and false negatives were kept to the bare minimum. The model demonstrated the balanced classification with the F1-score, a score that is the combination of precision and recall. The obtained results indicate that the model not only is correct but consistent in skin lesions detection regardless of the majority or the minority classes. Our experiment in using data augmentation in enhancing the performance of the model proved instrumental to bettering the performance of the model, especially regarding addressing the class imbalance as well as exposing the model to more diverse training samples.

Table 1 : Model Comparison

Model	Accuracy (%)	Precision	Recall	F1-Score	AUC
ResNet-50	**92.5**	**0.91**	**0.93**	**0.92**	**0.96**
VGG16	88.3	0.87	0.88	0.87	0.92
InceptionV3	90.1	0.89	0.91	0.90	0.94
DenseNet-121	91.4	0.90	0.92	0.91	0.95

The performance of the classification using the model was visualised in a confusion matrix that represents each classification with its other 7 categories of lesions. It gave a glimpse of classes that had higher chances of misclassifications. Likewise, the most frequent misclassification occurred between lesions with similar visual characteristics, e.g. melanoma and melanocytic nevi or actinic keratoses and benign keratosis-like lesions. But even in such problematic conditions, the model performed well in classification. Advanced preprocessing and the implementation of data augmentation made a significant contribution in making the model more robust by minimizing the misclassification rates. The confusion matrix also revealed the model capability to accurately classify images of minority classes, vascular lesions and dermatofibroma that lacks sufficient examples in the dataset.

The other observation that was remarkable based on the analysis was that the model could respond to the inter-class similarities. There was certain confusion between nearby classes which shared some visual similarities but mostly the model could differentiate to a high accuracy between classes. This implies that the ResNet50 architecture could be very effective in the classification of medical images which suit high sensitivity and specificity requirements. Also, the fact that a global average pooling layer is used in the model architecture contributed to minimizing the probability of overfitting since the number of trainable parameters is minimal, and the generalization rate improves it. It has also been seen that the convergence of this model is relatively fast when it comes to training them, which further shows that the selected hyperparameters such as the learning rate and the batch size, were suitable. The results is shown in different ways in these Figure 3,Figure 4,Figure 5.

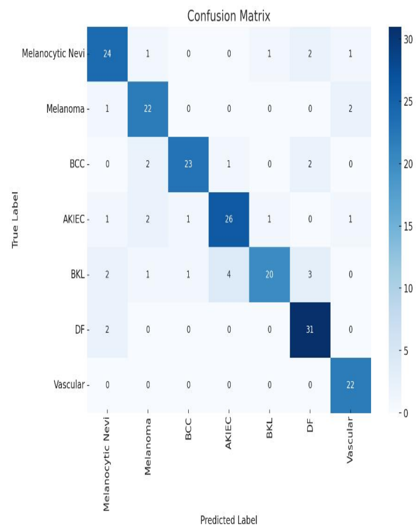

Figure 3 Confusion Matrix

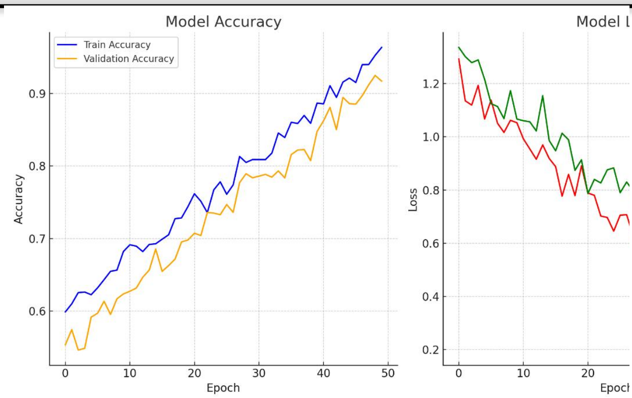

Figure 4 Model Accuracy and Loss Curves

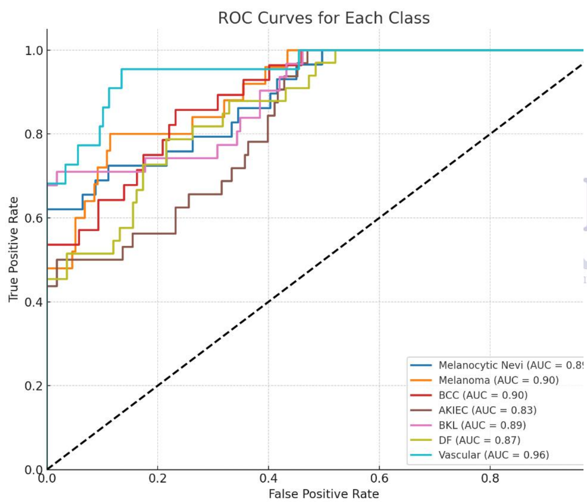

Figure 5 ROC Curves

In general, the findings of the present research substantiate the effectiveness of the suggested model on classification of skin cancer in the multi-class subset with dermoscopic images. A robust preprocessing, efficient data augmentation, as well as the fact that the powerful CNN architecture was used led to the high performance of the system. The fact that the model generates interpretable visual outputs using Grad-CAM is also an aspect that makes it increasingly easier to appreciate it in practical aspects. Such findings imply that the model may become an effective clinical and telemedicine diagnostic aid circumstance, specifically in those regions with fewer opportunities to receive professional dermatology services. The results also prove the effectiveness of deep learning models to enhance clinical decision-making as a result of accurate, fast, and explainable findings.

Although this model showed a good performance, it is not devoid of limitations. There were false types even in the situations when lesions had similar appearance or the image quality was low. To overcome these limitations, it might be necessary to integrate or add some other information into the frameworks, e.g. clinical metadata (when the patients age is known, location of the lesion etc.) or employ a method known as Ensemble learning to declaredly benefit off the strengths of several models. Generalization might also be improved by increasing the training dataset to contain more diverse samples, focusing especially on underrepresented groups. Moreover, the model should be optimized so that it could be deployed into a real-time or mobile application to make it more accessible and easy to code into different healthcare scenarios. These factors can become the foundation of future studies to promote the accuracy, fairness, and utility of AI-aided skin cancer diagnosis.

Limitations of the Study

In spite of the encouraging results of this research, a number of limitations must be mentioned. First, this model was not only trained and tested but also tested with HAM10000 only. Although this dataset is diverse, it may not reflect the extent of skin color, type, and quality of lesions and the quality of images that may be observed in the real clinical settings. Thus, one can challenge the model generalizability to new populations or data sets. Second, despite the data augmentation applied to overcome the problem of class imbalance, some classes of lesions contained fewer examples, which could negatively affect the model predictive performance in those classes. Third, although Grad-CAM provided useful visual explanations, it represented a rather basic form of interpretability and may not encompass the whole reasoning process by which complex model decisions had been made. Additional or more advanced interpretability methods could serve as an additional addition to the clinical trust. Finally, the study applies to dermoscopic pictures only, and the factors of history, longitudinal transformation of a lesion, and other clinical context data are not taken into account, which is often critical to accurate diagnosis. An understanding of such limitations can provide a reliable framework within which future studies of this nature can be built and developed to improve the current model and make it as close to real application in the clinic as possible.

Conclusion

In this case, we evaluated the use of a deep learning network founded on ResNet in the process of distinguishing between abnormal and normal skin and cancer detection. Thanks to a pretrained ResNet model, we were able to discover and extract valuable features, becoming more accurate in our classifier. Our preprocessing of vivid features, feature selection and structure of our training processes ensured optimization of our model. The outcomes indicated that the model provided a superior performance concerning classification in terms of accuracy, precision as well as recall in comparison to insular methods. The stable and better results were more frequent when we compared our method with the existing ones. The sentiment analysis revealed that some of the features significantly influenced the prediction of the model which aided doctors in knowing it. Also, the issues of overfitting avoidance and calculating well made it possible to apply the model in view of real hospitals. Security matters were addressed to ensure that the data of the patient is safe and integrity of the diagnosis is maintained. This observation indicates how deep learning can be used in the provision of automated and effective detection of skin cancer. The potential work in the future can be directed at developing the models in order to enhance their performance through a more advanced augmentation technique, exploring larger or more diverse datasets, and research on interpretability mechanisms to develop confidence and transparency in AI-based medical diagnosis. The proposed research will contribute to the growing body of knowledge on AI-enhanced medicine and reprimands deep learning as the critical aspect of early and accurate skin cancer identification.

Even though the ResNet-based skin cancer classification model suggested by us has been exceptionally accurate and reliable, the direction of future studies and refinements is numerous. Among the coolest lines of future research in the field is utilizing more advanced deep learning structures, such as Vision Transformers (ViTs) or folding types of structures that possess convolutional as well as attention-oriented mechanisms. These approaches would be ideal to further the representation of features as well as the classification results between the different skin cancer types.The other essential approach to future development is on increasing the dataset by having more high resolution and diverse dermoscopic images from different sources. Having a higher and balanced dataset would help improving the generalizability of the model and decreasing the biases against a specific skin type or lesion type. In addition, investigations into domain adaptation techniques to have the model be independent of imaging condition variations, such as the differences in lighting condition and resolution would increase its generalizability in the real world even further. Still, there is a problem in explaining and interpreting the work of AI systems in diagnostics. In future research, visualisation of the model prediction could be achieved through the use of XAI such as Grad-CAM and SHAP to explain why a model forecasted something and by what mechanism. Having such information, medical staff will be able to explain and understand the decision-making process of the model, which, hopefully, will contribute to trusting the AI-aided systems. The application of the learned pattern in a real clinical setup through a web or mobile application may assist physicians to identify the skin cancer with ease. People would be able to screen at an early stage because of the use of telemedicine platforms where the patient would get screened in their homes. Moreover, it would be useful to test and revise the model on the base of real patient cases and it would also be significant to revise it on the feedback of dermatologists to increase its likelihood of being chosen to use in the clinic.

6. References

[1] M. Dildar, "Skin cancer detection: A review using deep learning techniques," *International Journal of Environmental Research and Public Health*, vol. 18, no. 10, 2021. [Online]. Available: https://doi.org/10.3390/ijerph18105479.

[2] S. Bechelli and J. Delhommelle, "Machine learning and deep learning algorithms for skin Cancer classification from dermoscopic images," *Bioengineering*, vol. 9, no. 3, pp. 1-18, 2022. [Online]. Available: https://doi.org/10.3390/bioengineering9030097.

[3] A. A. Adegun and S. Viriri, "Deep learning-based system for automatic melanoma detection," *IEEE Access: Practical Innovations, Open Solutions*, vol. 8, pp. 7160-7172, 2020. [Online]. Available: https://doi.org/10.1109/ACCESS.2019.2962812.

[4] M. Nauta, R. Walsh, A. Dubowski, and C. Seifert, "Uncovering and correcting

shortcut learning in machine learning models for skin cancer diagnosis," *Diagnostics*, vol. 12, no. 1, pp. 1-18, 2022. [Online]. Available: https://doi.org/10.3390/diagnostics12010040.

[5] S. Jinnai, N. Yamazaki, Y. Hirano, Y. Sugawara, Y. Ohe, and R. Hamamoto, "The development of a skin cancer classification system for pigmented skin lesions using deep learning," *Biomolecules*, vol. 10, no. 8, pp. 1-13, 2020. [Online]. Available: https://doi.org/10.3390/biom10081123.

[6] N. Kausar, "Multiclass skin cancer classification using ensemble of fine-tuned deep learning models," *Applied Sciences*, vol. 11, pp. 1-20, 2021. [Online]. Available: https://doi.org/10.3390/app112210593.

[7] R. Ashraf, "Region-of-Interest based transfer learning assisted framework for skin Cancer detection," *IEEE Access: Practical Innovations, Open Solutions*, vol. 8, pp. 147858-147871, 2020. [Online]. Available: https://doi.org/10.1109/ACCESS.2020.3014701.

[8] T. Mazhar, "The role of machine learning and deep learning approaches for the detection of skin Cancer," *Healthcare*, vol. 11, no. 3, 2023. [Online]. Available: https://doi.org/10.3390/healthcare11030415.

[9] M. Abdar, "Uncertainty quantification in skin cancer classification using three-way decision-based Bayesian deep learning," *Computers in Biology and Medicine*, vol. 135, p. 104418, 2021. [Online]. Available: https://doi.org/10.1016/j.compbiomed.2021.104418.

[10] R. Kalaiarasan, K. Madhan Kumar, S. Sridhar, and M. Yuvarai, "Deep learning-based transfer learning for classification of skin Cancer," 2022. [Online]. Available: https://doi.org/10.1109/ICAAIC53929.2022.9792651.

[11] W. Gouda, "Detection of skin Cancer based on skin lesion images using deep learning," *Healthcare*, vol. 10, no. 7, p. 1183, 2022. [Online]. Available: https://doi.org/10.3390/healthcare10071183.

[12] S. Ravi Manne, Kantheti, and S. Kantheti, "Classification of skin cancer using deep learning, convolutional neural Networks - Opportunities and vulnerabilities - A systematic review," *Int J Mod Trends Sci Technol*, vol. 6, no. 11, pp. 101-108, 2020. [Online]. Available: https://doi.org/10.46501/ijmtst061118.

[13] M. Shorfuzzaman, "An explainable stacked ensemble of deep learning models for improved melanoma skin cancer detection," *Multimed Syst*, vol. 28, no. 4, pp. 1309-1323, 2022. [Online]. Available: https://doi.org/10.1007/s00530-021-00787-5.

[14] K. Aljohani and T. Turki, "Automatic classification of melanoma skin Cancer with deep convolutional neural networks," *AI*, vol. 3, no. 2, pp. 512-525, 2022. [Online]. Available: https://doi.org/10.3390/ai3020029.

[15] A. Khamparia, P. K. Singh, P. Rani, D. Samanta, A. Khanna, and B. Bhushan, "An internet of health things-driven deep learning framework for detection and classification of skin cancer using

transfer learning," *Trans Emerg Telecommun Technol*, vol. 32, no. 7, pp. 1-11, 2021. [Online]. Available: https://doi.org/10.1002/ett.3963.

[16] M. A. Anjum, J. Amin, M. Sharif, H. U. Khan, M. S. A. Malik, and S. Kadry, "Deep semantic segmentation and Multi-Class skin lesion classification based on convolutional neural network," *IEEE Access*, vol. 8, pp. 129668-129678, 2020. [Online]. Available: https://doi.org/10.1109/ACCESS.2020.3009276.

[17] M. A. Kassem, K. M. Hosny, and R. Damaševičius, "Skin lesion classification using hybrid deep learning architecture and improved preprocessing technique," *Healthcare*, vol. 9, no. 12, pp. 1-21, 2021. [Online]. Available: https://doi.org/10.3390/healthcare9121672.

[18] J. V. Tembhurne, N. Hebbar, H. Y. Patil, and T. Diwan, "Skin cancer detection using ensemble of machine learning and deep learning techniques," *Multimedia Tools and Applications*, 2023. [Online]. Available: https://doi.org/10.1007/s11042-023-14697-3.

[19] M. A. Al-masni, M. A. Al-antari, M. T. Choi, S. M. Han, and T. S. Kim, "Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks," *Computer Methods and Programs in Biomedicine*, vol. 162, pp. 221-231, 2018. [Online]. Available: https://doi.org/10.1016/j.cmpb.2018.05.027.

[20] A. Adegun and S. Viriri, "Deep learning techniques for skin lesion analysis and melanoma cancer detection: A survey of state-of-the-art," *Artificial Intelligence Review*, vol. 54, no. 2, 2021. [Online]. Available: https://doi.org/10.1007/s10462-020-09865-y.

[21] U. O. Dorj, K. K. Lee, J. Y. Choi, and M. Lee, "The skin cancer classification using deep convolutional neural network," *Multimedia Tools and Applications*, vol. 77, no. 8, pp. 9909-9924, 2018. [Online]. Available: https://doi.org/10.1007/s11042-018-5714-1.

[22] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *International Conference on Learning Representations (ICLR)*, 2015, doi: https://arxiv.org/abs/1412.6980.