Journal of Carcinogenesis
Covering Genesis to Therapeutics

Journal of Carcinogenesis, ISSN: 1477-3163
2025; Vol 24: Issue 2, https://carcinogenesis.com/

Deep Learning Based System for Early Detection and Classification of Melanoma Skin Cancer- A Systematic Review

Mrs. Komal A Champanerkar [1], Praniti Patil [2], Ms. Paurnima Prakash Patil[3], Chaitali Hitesh Mhatre[4], Vaishali A. Shirsath[5], Trupti Shah[6], Suresh Singh Rajpurohit[7], Dr.Swastika [8]

[1]Assistant Professor, Thakur Shree DPS college of Engineering and Management, Maharashtra 401202,

Email ID: champanerkar.komal@gmail.com

[2]Assistant Professor, SVKM's Dwarkadas J. Sanghvi College of Engineering, Mumbai, Maharashtra 400056,

Email ID: pranitirpatil@gmail.com

[3]Assistant Professor, South Indian Education Society Graduate School of Technology, Nerul, Mumbai India,

Email ID: paurnimap@sies.edu.in

[4]Assistant Professor, Universal College of Engineering, Vasai, Maharashtra 421302, India

Email ID: chaitali.mhatre@universal.edu.in

[5]Assistant professor, Vidyavaridhi's College of Engineering and Technology, Maharashtra 401202, India

Email ID: .vaishali.shirsath@vcet.edu.in

[6]Assistant Professor, Thakur college of Engineering and Technology, Mumbai, Maharashtra 400101, India

Email ID: .trupti.shah@tcetmumbai.in

[7]Suresh Singh Rajpurohit, Assistant Professor, Thakur Shyamnarayan Engineering College (TSEC),[redacted]Mumbai, Maharashtra 400101,

Email ID :Srajpurohit422@gmail.com

[8]Assistant Professor, R P Sharma Institute Of Technology,[redacted]Patna, Bihar 801503,

Email ID: Swastikaanuj54@ gmail.com

ABSTRACT

Melanoma is the most dangerous form of skin cancer and its detection at the initial stage is very important to increase patient's survival. This paper compares the use of deep learning systems in the early identification and categorization of melanoma skin cancer. The research used a melanoma cancer dataset obtained from Kaggle which had 9600 training images, and 1000 test images divided equally between the benign and malignant classes. The pre-processed datasets are employed to train transfer learning models, including VGG16, Mobile Net, Dense Net, ResNet50 and InceptionV3. Model performance is assessed by accuracy, loss, Cohen's kappa, precision, recall, f1-score and confusion matrix. The best performing model is shown to be deployable using Streamlit, which provides a user interface through an input image, on-the-fly pre-processing, and real-time classification. The proposed system is better than the other methods and can be a good and practical approach for melanoma detection..

Keywords: *Melanoma detection, Transfer learning, Deep Learning, Image Processing, VGG16, MobileNet, DenseNet, ResNet50, InceptionV3.*

How to Cite: Mrs. Komal A Champanerkar , Praniti Patil , Ms. Paurnima Prakash Patil, Chaitali Hitesh Mhatre, Vaishali A. Shirsath, Trupti Shah, Suresh Singh Rajpurohit, Dr.Swastika , (2025) Deep Learning Based System for Early Detection and Classification of Melanoma Skin Cancer- A Systematic Review, *Journal of Carcinogenesis*, *Journal of Carcinogenesis*, *Vol.24, No.2*, 26-34

1. INTRODUCTION

Melanoma is the most aggressive and life-threatening form of skin cancer that is prevalent in the world today. It is a more dangerous form of skin cancer than the others because it is more likely to metastasis and cause severe health complications if not detected early. The World Health Organization (WHO) estimates that more than 300,000 new melanoma cases are diagnosed every year and the incidence rates are increasing worldwide. These methods are critical to improving survival rates if detected and treated early; however, traditional diagnostic methods rely on subjective interpretation by dermatologists, which can result in variable results.

The use of medical imaging and artificial intelligence has changed the detection and classification of skin cancer. Deep learning, a form of AI, has been found to perform very well in a variety of medical diagnostics that involve images. CNN, a type of deep learning architecture, has been found to perform especially well in the analysis of thermoscopic images for melanoma detection. CNNs work by learning features directly from raw images, thus not requiring manual feature extraction to improve the efficiency and accuracy of the diagnostic process.

However, there are some challenges that remain in application of deep learning in melanoma detection. The training datasets are also quite sensitive to the quality of the images, lighting and skin tones and these complications make the detection task even harder. To overcome these challenges, it is crucial to have good pre-processing methods, augmentation techniques and suitable model configurations to enhance the robustness and accuracy of the detection model.

Another important issue is the understandability of deep learning models in medical environments. For using the model in clinical practice, the physician must understand the basis of the model's suggestions. Techniques such as highlighting regions of interest in the dermoscopic images and have been used to explain the model decisions visually. Such interpretability is not only important for the acceptance of AI-based tools in clinical practice but also helps the clinicians to work in tandem with the AI systems.

The purpose of this study is to conduct a systematic review of the recent advances in deep learning-based frameworks for melanoma detection and differentiation. It assesses the qualities of the existing models in terms of accuracy, loss, Cohen's kappa, precision, recall, f1 score and confusion matrices to determine failure factors and where best efforts can be directed. In addition, the study suggests an improved strategy that incorporates knowledge transfer, combination strategies, and sophisticated data preparation to overcome the deficiencies of the current approaches.

The main purpose of this research is to design an effective, efficient and understandable system that may support dermatologists in identifying melanoma at its earliest stage. Thus, the proposed system combines the advantages of deep learning and solves its problems, so it has the potential to decrease the misdiagnoses rate, improve the patients' health and the AI integration in healthcare. In this paper, we would like to present our contribution to the fight against melanoma and improvement of dermatological practice because of this study.

2. LITERATURE REVIEW

Detecting skin cancer at an early stage is very important as it can help to increase the chances of successful treatment and improve the survival rates. Recently, there has been a great improvement in the automated skin cancer detection systems by machine learning (ML) and deep learning (DL) techniques. This review includes the major studies that have been carried out within this area, and it examines the different approaches, sources of data, and results measurement tools.

Study	Methodology	Dataset	Key Findings	Accuracy
Skin Cancer Facts & Statistics. Skin Cancer Foundation	Histogram-based local descriptors (LBP, LDN, PHOG, LDiP, MBC) with SVM and XGBoost classifiers	Dataset 1 from Kaggle, HAM10000	Proposed color histogram-based features improved classification success.	70% (Dataset 1), 76.50% (HAM10000)

Study	Methodology	Dataset	Key Findings	Accuracy
Winkler et al.	Convolutional Neural Networks (CNN)	Dermoscopic images	CNN model achieved dermatologist-level accuracy in melanoma detection.	Not specified
Brinker et al.	Deep Learning (CNN)	Dermoscopic images	Deep learning surpassed human dermatologists in melanoma classification tasks.	Not specified
Munir et acccl. (2021)(Radha et al.,	Deep Neural Networks	Dermoscopic images	Achieved dermatologist-level classification of melanoma.	Not specified
Sood et al.	CNN	Dermoscopic images	Demonstrated promise of CNNs in enhancing skin cancer diagnosis accuracy.	Not specified
Hekler et al.	Deep Learning	Histopathological images	Deep learning models outperformed pathologists in melanoma classification.	Not specified
Auxilia Osvin Nancy et al. (2023)(Nancy et al., 2023, p. 2)	Various ML and DL algorithms	ISIC archive	RF outperformed other ML algorithms; MobileNetv2 and ensemble models showed superior performance.	77.58% (with augmentation)
Harsh Bhatt et al. (2022)(Bhatt et al., 2023, p. 1)	KNN, SVM, CNN	Various datasets	Reviewed state-of-the-art techniques; highlighted the effectiveness of CNNs for skin cancer detection.	Not specified

Study	Methodology	Dataset	Key Findings	Accuracy
Rifi et al. (2019)(Bhatt et al., 2023, p. 7)	KNN, ANN	ISIC datasets	Achieved high accuracy in skin lesion identification using hybrid techniques.	88.84% (KNN), 87.8% (SVM)
Elgamal (2018)(Bhatt et al., 2023, p. 7)	Hybrid technique (DWT, PCA, KNN, ANN)	40 images	Developed a CAD system for skin cancer classification with promising results.	88.84% (KNN), 85% (ANN)

3. PROPOSED METHODOLOGY

The proposed research is aimed at the development of an advanced deep learning-based system for the early detection and classification of melanoma skin cancer. Melanoma is a highly aggressive form of skin cancer, and it needs proper and accurate diagnosis to enhance the patient's prognosis. This study is meant to contribute to the solution of the key issues in melanoma detection, such as the problem of dermoscopic images variability, image pre-processing and the requirement for reliable interpretability. The methodology is planned to develop a robust and clinically applicable diagnostic tool by leveraging state-of-the-art transfer learning techniques. The system incorporates pre-processing steps, optimal model architectures and rigorous evaluation metrics to achieve high performance and generalization across various datasets.

The approach starts with data pre-processing, where dermoscopic images are first Resized to 224×224 pixels, then Converted to grayscale version of the image, then Noise reduction using block- matching 3D (BM3D) filtering, then Binary mass lesion image and Binary lesion mask formed by thresholding, then Segmentation using binary lesion masks and morphological operations, then Generation of masked ROI showing only the masked area of the image and inverse-masked images showing everything except the masked area, then Canny edge-detected image to enhance model training.

For transfer learning based pre-trained architectures, we need to give images of a certain size. This ensures that there is uniformity in input dimensions during batch processing during training and inference. This step also reduces computational complexity by down sampling high resolution images whilst keeping important features.

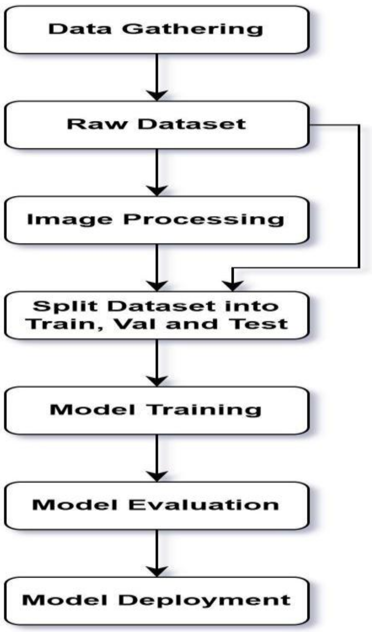

Fig. 1. Proposed Methodology Architecture

Dermoscopic images are usually in RGB format and are therefore redundant in color information. However, for melanoma detection, intensity variations and texture patterns are more relevant than color. Grayscale conversion of RGB images reduces the data to a single channel, keeping only the luminance (intensity) values. This simplifies the input of the model without jeopardizing the necessary information for the lesion analysis. The weights (0.2989, 0.5870, and 0.1140) are based on human perception, with green contributing the most to brightness, followed by red and blue.

In this case, we employ several preprocessing steps that are crucial to model training to ensure equality, effectiveness, and increased reliability of the model. This is because different images have different sizes which may cause batch learning to be inefficient and vary in some cases. The grayscale conversion simplifies the feature space by reducing the number of elements in the input data without sacrificing relevant information that is helpful for the lesion detection task. In a similar manner, these steps help to improve the dataset to be compatible with deep learning by reducing memory requirements and the computational time. Furthermore, the standardized and simplified inputs allow the model to learn more effective patterns and improve the generalization performance of the model.

Block-Matching 3D (BM3D) filtering is a typical method for filtering noise and improving the quality of images by grouping similar patches of the image in 3D blocks and filtering them jointly. This method is divided into two parts: initial and final denoising. In the basic step, the patches are organized according to similarity and transformed into the 3D DCT domain. A threshold is applied to remove noise, and the inverse transform is used to rebuild the denoised patches. In the final step, more sophisticated refinement is performed using information fusion from overlapping blocks.

BM3D filtering is crucial for enhancing model training by removing noise without compromising on important image structures like lesion boundaries. This denoising ensures that the input images are devoid of random variations or artifacts that might confuse the model. In addition, the focus on clean high-quality features, by the BM3D filtering, enhances the model's capacity to learn patterns significant to melanoma detection, including texture and shape. Furthermore, the lowered noise leads to more consistent feature representations that enhance learning, reduce convergence time and increase the generalization capability on the unseen data.

Using the original image, binary mass lesion images and lesion masks are created by thresholding the image to separate the lesion region from the background. In the case of thresholding, grayscale pixel values are converted to binary values (0, 1) based on a threshold. Should the intensity of a pixel be greater than the threshold, it is deemed to be part of the lesion (foreground) whereas if it is below the threshold, it is part of the background. This produces a binary mass lesion image that outlines the region of the lesion, and the corresponding lesion mask is a template for extracting or focusing on the lesion in subsequent processing steps.

The generation of binary mass lesion images and lesion masks helps to improve the training process by data simplification, by separating the lesion from the skin background. This focused region of interest (ROI) avoids the need for the model to learn irrelevant background information, and instead it focuses on the structure and texture of the lesion. By reducing the level of complexity and focusing on the relevant parts, these binary images enhance the accuracy of the segmentation and assist the model to learn important patterns faster. This in turn results in better generalization, faster convergence of training and better performance in melanoma detection.

For segmentation, the region of interest (ROI) i.e. the lesion area is isolated from the rest of the image using a binary mask. The binary lesion mask renders the lesion white (value 1) and the rest of the image black (value 0). Dilatation, Erosion, Opening, and Closing are the morphological operations that are applied to fine tune the segmentation to eliminate small noise or to fill up gaps in the lesion region. Without using a structuring element, these operations have a poor capability of producing a good lesion segmentation.

Using binary lesion masks and morphological operations in the training process ensures that the model does not consider the information relevant for analysis other than the lesion area. This precise segmentation helps the model to concentrate on the important features like shape, texture and boundaries of the lesion which in turn enhance the model's capability to discriminate between melanoma and non-melanoma lesions. In addition, the morphological operations are used to improve the mask by erasing small objects and outliers that might mislead the model. Hence, by giving the model cleaner and more accurate input data, it can learn better, resulting in enhanced segmentation performance and better accuracy in detecting melanoma.

This applies the binary lesion mask to separate out and emphasize the lesion region as the Region of Interest (ROI) and to generate an inverse mask to exclude the lesion. The masked ROI has the lesion region only and the rest of the regions are set to black (value 0). The opposite is the case with the inverse masked image where the non-lesion regions are kept, and the lesion region is set to black.

The generation of masked ROI and inverse-masked images enhances model training by providing two types of data: one focused solely on the lesion area (masked ROI) and the other on the surrounding background (inverse-masked image). This dual approach helps the model learn to recognize both the presence of the lesion and distinguish it from the background. The isolated lesion in the masked ROI allows the model to focus on the relevant features, improving classification accuracy. In contrast, the inverse-masked image can be used to emphasize background features, improving the model's ability to differentiate between the lesion and non- lesion regions, ultimately leading to more accurate and robust melanoma detection.

The Canny edge detection algorithm is used to detect edges in an image and is particularly useful for finding the edges of lesions. The algorithm works by applying a series of steps: First, it blurs the image using a Gaussian filter to ignore noise; second, it calculates gradient magnitude and direction to find edges; third, it applies non-maximum suppression to thin edges; lastly, it performs edge tracking by hysteresis to keep the strong edges and eliminate the weak ones.

Canny edge detection enhances model training by focusing on the boundaries of the lesions, which are often key in correct classification. Edge features are thus highlighted by the model to better distinguish between lesion and non-lesion regions, particularly when textural or intensity changes are subtle. This pre-processing step enables the model to learn key structural features such as shape and contours of melanoma lesions, thus enhancing its decision-making capability. Furthermore, edge-detection reduces the complexity of the image by highlighting important features, thereby facilitating efficient identification and classification of lesions by the model.

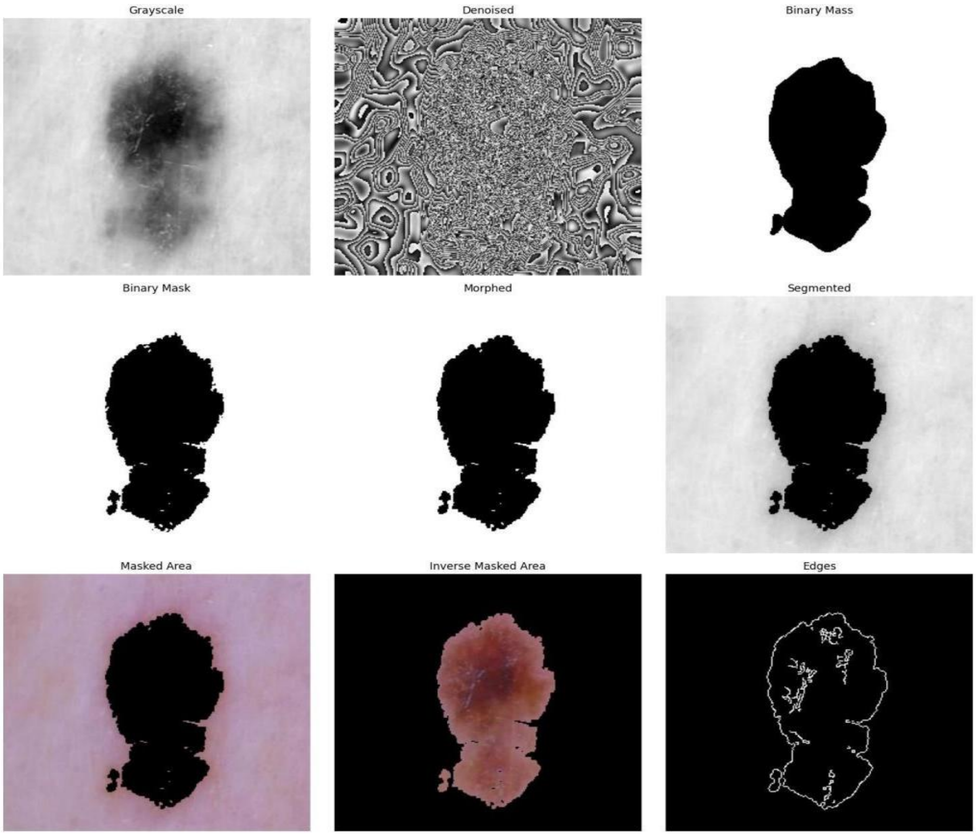

Fig. 2. Image Pre-processing

First, in the dataset, through image pre-processing, essential features such as the lesion boundaries and the region of interest (ROI) are enhanced and isolated. Then, Model Selection and Training follows. In this paper, advanced transfer learning models, including VGG16, MobileNet, DenseNet, ResNet50 and InceptionV3 are employed for melanoma classification. These pre-trained models are fine-tuned to fit the specificities of the melanoma dataset. Fine-tuning is the process of continuing training of the pre-trained models, i.e., adjusting their weights to learn domain-specific features while retaining the generalizable patterns learnt during pre-training. Both raw and pre-processed images are used in training these models to enable a comparison of the performance of the datasets. This approach takes advantage of the models' powerful feature extraction capabilities and focuses on structural, textural, and boundary features in the images in the effort to accurately detect and classify melanoma lesions between malignant and benign cases. In the case of transfer learning models, the performance was evaluated using both raw and pre-processed datasets, and accuracy, loss, Cohen's kappa, precision, recall, and F1-score were used as metrics in the evaluation, as shown in Table 1 and Table 2. On the raw dataset, DenseNet provided the best accuracy of 90.98%, while MobileNet was close behind at 89.31%. However, when the pre-processed dataset was used, the performance of all the models increased remarkably. With the pre-processed dataset, MobileNet was the best performing model with an accuracy of 91.95%, a loss of 8.05% and a Cohen's kappa score of 83.86%. This shows that there is better agreement between the actual and predicted classifications. Moreover, MobileNet had a precision, recall and F1-score of 92, which shows that it is less likely to make a wrong decision in a melanoma classification task. The above tables show that the value of all the metrics increases after pre-processing, which shows that pre-processing helps the model learn better features. Hence, MobileNet appears to be the best model for this task among the models considered.

Table 1. Classification report with raw dataset

With Raw Dataset	Accuracy	Loss	Cohens Kappa	Precision	Recall	F1 Score
VGG16	87.65	12.35	75.31	88	88	88
Mobilenet	89.31	10.69	78.6	89	89	89
ResNet50	79.94	20.05	60.06	80	80	80
Dencenet	**90.98**	**9.02**	**81.92**	**91**	**91**	**91**
Inceptionv3	88.48	11.52	76.93	88	88	88

Table 2. Classification report with pre-process dataset

With Pre- process Dataset	Accuracy	Loss	Cohens Kappa	Precision	Recall	F1 Score
VGG16	91.33	8.67	82.59	91	91	91
Mobilenet	**91.95**	**8.05**	**83.86**	**92**	**92**	**92**
ResNet50	82.72	17.28	65.16	84	82	82
Dencenet	91.6	8.4	83.17	92	92	92
Inceptionv3	89.45	10.55	78.87	89	89	89

4. RESULTS

The results of the experiments showed that the use of preprocessed datasets enhanced the performance of the model. Of the five transfer learning models that were evaluated in this study, VGG16, MobileNet, DenseNet, ResNet50 and InceptionV3, MobileNet was the most effective, achieving 91.95% accuracy, 8.05% loss and the highest Cohen's kappa score of 83.86% on the pre-processed dataset. These metrics show that MobileNet is more effective at identifying melanoma lesions and is more robust to the complexities of dermoscopic images. The results also revealed that preprocessing was important because all the models improved their precision, recall and F1-score with the use of preprocessed data, and MobileNet had values of 92% for all of these metrics, which indicates consistent performance.

The comparison between raw and pre-processed datasets revealed the critical role of pre- processing in enhancing model efficacy. While DenseNet showed strong performance on the raw dataset with an accuracy of 90.98%, MobileNet surpassed it with pre-processed data, demonstrating the effectiveness of leveraging edge detection, lesion segmentation, and ROI masking for feature enhancement. These steps reduced noise and background distractions, enabling the models to focus on relevant lesion features. The improved metrics also validated the hypothesis that carefully curated and pre-processed data can significantly enhance model generalization and predictive capabilities.

To realize the real-time melanoma detection application, Streamlit was used to develop the user interface to receive input from the models. The pipeline defined for preprocessing is used to preprocess the dermoscopic images that the users upload. In addition, MobileNet, the best performing model, is chosen to classify the lesions as benign or malignant. The deployment does not only predict but also visualize the preprocessed images including masked ROI and edge detected versions of the images. This capability is useful for clinicians and researchers to understand the characteristics of the input data that may influence the model's decision.

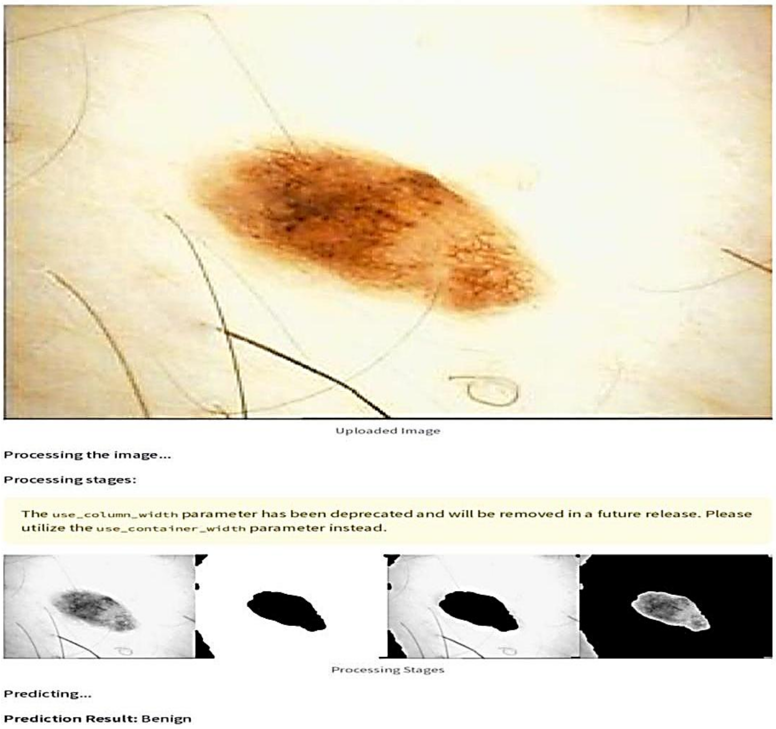

Fig. 3. Deploy the trained model

Using Streamlit for the deployment guarantees the solution's usability and scalability, so that it can be easily incorporated into clinical practice. Its light-weight architecture is compatible with cloud platforms, which supports its use by many without the need for much infrastructure. The application of the latest AI methods combined with an easy-to-use interface shows the future of the technology within the healthcare setting, providing a much needed and effective tool for the early detection and diagnosis of melanoma. This transition from research to application is thus significant from a practical perspective, as it bridges the gap between computational advances and actual medical problems.

5. CONCLUSION

A comprehensive approach to the detection of melanoma skin cancer was introduced, with the use of sophisticated transfer learning models in conjunction with a strong data pre-processing. The models' performance was enhanced by increasing the dataset using techniques like resizing, grayscale conversion, noise reduction, and lesion masking. MobileNet was identified as the most efficient model. It had the best accuracy of 91.95% and was the best generalizing model across all the evaluation criteria. These techniques' integration reveals the significance of data preparation and model choice in producing reliable and accurate diagnostic systems for vital healthcare applications. Moreover, the application of the model was demonstrated through Streamlit, which showed the feasibility of the model in real-time detection for clinicians and researchers. The proposed framework can, therefore, be used as a reference for the development of similar models for other skin conditions. Future work should also focus on enhancing the model's robustness and generalization capabilities across different demographics and clinical settings by integrating larger and more diverse datasets. Moreover, the incorporation of multi-modal data including patient history and genetic data may improve the system's decision making. The integration of XAI techniques to explain the model's decision-making process will increase the model's credibility in clinical practice. Extending the deployment to mobile and cloud platforms will help make this technology available for remote and underserved areas, thus increasing its potential in solving global healthcare issues. Thus, the model trained can be deployed using Streamlit for it to be easily applicable in practice.

REFERENCES

[1] Skin Cancer Facts & Statistics. Skin Cancer Foundation. Accessed January 3, 2024.

[2] Rogers HW, Weinstock MA, Feldman SR, Coldiron BM. Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the U.S. population, 2012. JAMA Dermatol.

[3] Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks.

[4] Young AT, Fernandez K, Pfau J, Reddy R, Cao NA, von Franque MY, et al. Stress testing reveals gaps in

clinic readiness of image-based diagnostic artificial intelligence models. NPJ Digit Med.

[5] Liu Y, Jain A, Eng C, Way DH, Lee K, Bui P, et al. A deep learning system for differential diagnosis of skin diseases. Nat Med.

[6] Haenssle HA, Fink C, Schneiderbauer R, et al. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists.

[7] Brinker TJ, Hekler A, Enk AH. Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification.

[8] Winkler JK, Schäfer I, Bender C. Diagnostic performance of a Convolutional neural network for dermoscopic melanoma recognition in clinical practice. Journal of Investigative Dermatology.

[9] Aydin Y. A Comparative Analysis of Skin Cancer Detection Applications Using Histogram-Based Local Descriptors. Diagnostics.

[10] Auxilia Osvin Nancy, Prabhavathy P, Arya Meenakshi S, Ahamed B. Comparative study and analysis on skin cancer detection using machine learning and deep learning algorithms. Multimedia Tools and Applications.

[11] Wei ML, Tada M, So A, Torres R. Artificial intelligence and skin cancer. Frontiers in Medicine.

[12] Bhatt H, Shah V, Shah K, Shah R, Shah M. State-of-the-art machine learning techniques for melanoma skin cancer detection and classification: a comprehensive review.

[13] Dildar M, Akram S, Irfan M, et al. Skin cancer detection: a review using deep learning techniques. Int J Environ Res Public Health.

[14] Alquran H, Qasmieh IA, Alqudah AM, et al. The melanoma skin cancer detection and deep learning methods for cancer detection. J Image Process AI.

[15] Daghrir J, Tlig L, Bouchouicha M, et al. Melanoma skin cancer detection using deep learning and classical machine learning techniques