State of the art deep learning methods for melanoma classification

Niharika Thakur[1]

{niharika.thakur@chitkara.edu.in[1]}

Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab[1]

Abstract. Melanoma is a skin cancer, which spreads at the faster rate with higher motality if left untreated. Latest advancements in deep learning (DL) are changing the way of melanoma diagnosis employing traditional clinical and dermoscopic investigations performed by medical experts. Emphasising on the convolutional neural networks (CNNs), and its utilization in the field of melanoma diagnosis, this study examines state-of-the-art approaches used till date. By using the large annotated image samples, CNNs can capture subtle patterns in shape, texture and color of lesions, resulting in comparable accuracy as that of dermatologists. Also, the performance of diagnosis could be further improved by models like pre-trained transfer learning. But, issues of restricted demographic validation, interpretability of models, and possible biases against people having darker tones still exists. Although, there are still gaps in the use of more recent designs such as transformers, few recent work include merging of segmentation and classification after pre-processing to create end-to-end pipelines. Broad deployment requires addressing clinical integration, data diversity, and resilience. Building on these developments, DL can improvise early diagnosis of melanoma, lower mortality and reduce patients across a range of demographics.

Keywords: Melanoma; Classification; Deep learning; Dermoscopy images.

1 Introduction

The most deadliest form of skin cancer, melanoma, is caused by melanocytes growing out of control. Despite being less frequent than other skin cancers, melanoma spreads quickly and, if left untreated, can have a high death rates. Early diagnosis is essential since advanced forms of melanoma frequently have a poor prognosis [1]. Dermatologists often use clinical examinations and dermoscopic investigations for diagnosis, but it highly depends on the expertise of doctor. Current research indicates that the deep learning (DL) approaches, commonly used in medical imaging offers improvised accuracy for diagnosis of melanoma, thereby increasing liklihood of earlier diagnostics [2]. Studies show that diagnosing melanoma timely resulted in survival rates above 90% with its advancements globally. Traditional methods of diagnostics i.e. ABCDE (Asymmetry, Border, Color, Diameter, Evolving) relied majorly on expertise of dermatologists, which might be prone to certain human errors [3]. Thus, convovolutional neural networks (CNNs), one of DL is found to be useful diagnostic tool for melanoma due to its self learning and automated nature of diagnosis. Furthermore, it also provides abitlity to pre-process, segment and classify lesions as malignant or beningn [4]. Larger imaging samples having annotations, facilitates training of models in varied instances offering success of deep learning models. In

IICTDS 2025, December 19-20, Chandigarh, India
Copyright © 2026 EAI
DOI 10.4108/eai.19-12-2025.2362582

case of lesions classified as benign and malignant, pre-trained algorithms can even surpass performance of dermalogists in terms of efficiency and time [5]. Significant advances have been done in recent years due to annual ISIC challenge featuring categorization of skin lesions [6]. Varied networks like ResNet, InceptionNet and VGGNet have been employed, which offered increased flexibility of model solving issues such as vanishing gradient, enabling efficient networks [7]. In melanoma classification, transfer learning is also widely used since it enable models that have already been trained on generic datasets, such as ImageNet, to be optimized for medical pictures, lowering the need for labelled data and speeding up training [8]. Deep learning holds promise, but there are still obstacles to overcome. There is a chance that melanoma lesions may be misclassified due to their wide range of appearance, especially in the early stages. Furthermore, the lack of transparency in deep learning models makes it difficult for clinicians to trust system findings. Model biases create ethical problems since they run the danger of under diagnosing individuals with darker complexion, especially when they favour lighter skin tones due to imbalanced datasets [9]. In an effort to overcome these gaps and create equitable models across a range of populations, researchers are concentrating more on fairness aware algorithms. Widespread clinical use depends on ongoing advancements in dataset variety, and model interpretability. The ability of deep learning to diagnose melanoma early might greatly increase survival rates and enhance medical results.

2 Related Work

Few studies performed in this field till date are as follows:

Majtner et al. [1] used modified CNN instead of neural networks to categorize their data for extraction of features and pre-processing. This technique used 900 training samples from the International Skin Imaging Collaboration (ISIC), an open-access dataset. The imaging samples were classified into malignant and benign cases. During pre-processing, the image size was downsampled and grayscaled. After that, the system used AlexNet to extract features by using binary masking and bounding box approaches. The accuracy of the suggested approach was found to be 86% with 99.9% specificity.

A two-stage approach was proposed by Vipin et al. [3], which involved segmentation and classification. Reducing human error and time in melanoma detection procedure is the main objective of developing such system. The system employed a dataset of 13,000 images of which 7,353 were left after elimination of non-usable images. The segmentation phase made use of the symmetric U-Net, which is composed of three main parts: Upsampling, Downsampling and Bottleneck. The recurrent neural networks and CNN were employed to train the classifier. With the loss function as binary cross-entropy, 91% recall and 88.7% accuracy was attained.

Nasr Esfahani et al. [2] suggested modified CNN to fetch significant details for categorization after pre-processing. This study made use of 170 non-dermoscopic images. For improved training, the dataset was expanded to include 6,120 original and generated images. Gaussian filtering, mask creation, and lightning correction were used for pre-processing. One input layer, two convolutional and max-pooling layers, each placed alternately, and a final Fully connected (FC) layer made the CNN architecture. The findings were divided into two groups: nevus and melanoma. The accuracy and specificity of the advised approach were 81% and 80%, respectively.

Attia et al. employed encoding and decoding based CNN topologies namely Seg-Net and FCN [9]. Seven convolutions along with two max pool layering were used in the suggested architecture to encode input images. The performance of the approach was improvised by combining recurrent networks with different memory architectures. The competition, "Skin lesion analysis towards melanoma detection" in 2016 provided lesion images for model training and testing. The accuracy and specificity of the suggested model were 98% and 94% respectively.

Mukherjee et al. [7] suggested the convolutional architecture for detection of malignant lesions. The model components were represented by three two-dimensional (2D) convolution, two max-pool, one input layer and a FC layering at the terminal. Prediction was done by soft-max layering, Dermofit (13,000 images) and MEDNODE (1,700 images) were the two datasets that were integrated. For each dataset, accuracy rates were 90.14% and 90.58%, respectively. On merging of datasets, accuracy achieved was 83.07%. Compared to previous studies and conventional models, the performance of the suggested model was found to be superior.

Using two convolutional, two max pool, and a FC layer, Sanketh et al. [8] suggested CNN based model. The model's primary function was to recognise skin cancer timely, so that treatment could begin and death rate could be decreased. ISIC provided the dataset, which included 2,719 images in total that were divided into two groups: benign and malignant. Here, 816 testing images and 1,906 training images were utilized for experimentation. The ideal outcome of 98% was attained, with number of factors and their varied values.

AlexNet and VGG16 pertained networks were used in two distinct ways by Gulati et al. [5] as a paradigm for extraction of features and transfer learning. The Dermatology Service at Hospital Pedro Hispano in Portugal provided the PH2 dataset. There were 200 images in total, 160 of which fall into the benign and rest into malignant group. The required pre-processing was carried out since the aforementioned networks needed inputs of varying sizes. Extraction of features and transfer learning was done by each VGG16 and AlexNet separately. Here, VGG16 outperformed the other approaches with a specificity of 96.87% and an accuracy of 97.5%.

In the hybrid approach by Daghir et al. [10], conventional techniques like KNN and SVM were combined with CNN. The dataset having 640 images were used, out of which 128 were selected for testing and 512 for training. After the input layer, the CNN having nine layers including two dropout layers, three convolutional layers, three max pooling layers and last FC layering was used in analysis. Pre-processing methods used here were morphology, snakes and the Otsu method. Here, the suggested CNN method gave improved accuracy of 85.5% as compared to other methods.

Acousta et al. [6] used a pretrained ResNet structure in conjunction with a CNN approach that was both mask-based and region-based. There were two phases to the architecture's implementation. The lesions's border box was made in the initial step using Mask RCNN, and ResNet in second staging was employed for categorizing the lesion as either benign or malignant. With an accuracy of 90.4% and specificity of 92.5%, the suggested model outperformed the top performances ISIC 2017 models by 3.66%.

Using the Keras Sequential API, Rahi et al. [4] suggested a network and contrasted its outcomes with networks like DENSENET121, RESNET50 AND VGG16. ISIC repository comprising of 10,015 images were used for experimentation. Four convolutions, two max pool, and one fully

connected layer for experimentation were suggested by 6-layer CNN model. The accuracy of the model was 79%. VGG16, RESNET, and DENSENET were then used on the same dataset; where DENSENET gave highest accuracy of 90%.

Further, Al-Rasheed et al. [11] and Rriya Natha et al. [12] suggested Ensembling based approach for diagnosis offering Accuracy of 93.5% and Recall of 95.80% respectively.

Recently, Alotaibi & AlSaeed [13] employed Xception network offering 94.11% accuracy for classifying dermoscopic images into benign and malignant cases.

Table 1 shows the comparison of approaches discussed above

Table 1. Comparative analysis

Authors	Approach Used	Performance
Majtner et al. [1]	Alexnet	Accuracy: 86% Specificity: 99.9%
Vipin et al. [3]	U-net and Recurrent neural network	Accuracy: 88.7% Recall: 91%
Nasr Esfahani et al. [2]	Convolutional neural network	Accuracy: 81% Specificity: 80%
Attia et al. [9]	SegNet and Fully connected network	Accuracy: 98% Specificity: 94%
Mukherjee et al. [7]	Convolutional neural network	Accuarcy: 83.07%
Sanketh et al. [8]	Convoltional neural network	Accuracy: 98%
Gulati et al. [5]	AlexNet and VGG16	Accuracy: 97.5% Specificity: 96.87%
Daghir et al. [10]	Convolutional neural network	Accuracy: 85.5%
Acousta et al. [6]	Mask Recurrent neural network and ResNet	Accuracy: 90.4% Specificity: 92.5%
Rahi et al. [4]	VGG16, ResNet, DenseNet	Accuracy: 90%
Al-Rasheed et al. [11]	Ensemble	Accuracy: 93.5%
Priya Natha et al. [12]	Max voting ensemble	Recall: 95.80%
Alotaibi & AlSaeed [13]	Xception	Accuracy: 94.11%

3 Gaps in the existing studies

The following are some potential limitations and gaps in the diagnosis of melanoma based on studies till date.

- Although deep learning features are optimized in recent study, feature interpretability might aid physicians in understanding and trusting models. Furthermore, there is no discussion of data augmentation and its effect on robustness.
- Conventional CNN architectures still require improvements. Also, it is not validated on a variety of demographic data, which could affect how broadly the findings can be applied.
- The integration of segmentation data with classification models to establish an end-to-end diagnostic pipeline is not examined in this study.

- The findings of this study might not be very applicable to image-based melanoma diagnosis because text and images are from different data types.
- More intricate designs have shown promising results in classification of medical images, such as transformer-based models or ensemble approaches, might not have been explored in this study.
- Fairness factors such as possible biases in datasets may affect the accuracy of melanoma diagnosis across various ethnic groups.

Overcoming these gaps could improvise the performance and equity of melanoma diagnosis methods in clinical settings.

4 Conclusion

In conclusion, the precise and reliable identification of melanoma from dermoscopic images has shown great potential to models like improved CNNs. The advent of networks like ResNet, InceptionNet and VGGNet has made it possible to differentiate amongst malignant and benign cases. Transparency in model, data biases and limited generalizations between varied populations are issues existing despite of these developments. Automated diagnosis of melanoma need to be fair and dependable in clinical settings to resolve the problems. Robustness in the models might be increased and diagnostic accuracies could be resolved with hybrid classification and segmentation, thereby increasing demographic variety of datasets. Performance could also be improved by investigations of intricate structures like transformer based and hybrid models.

References

[1] Majtner, T., Yildirim-Yayilgan, S., Hardeberg, J.Y.: Optimised deep learning features for improved melanoma detection. Multimedia Tools and Applications. 78(9), pp. 11883-118903 (2018)
[2] Nasr-Esfahani, E., Samavi, S., Karimi, N., et al.: Melanoma detection by analysis of clinical images using convolutional neural network. In: 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE (2016)
[3] Vipin, V., Nath, M.K., Sreejith, V., et al.: Detection of melanoma using deep learning techniques: A review. In: 2021 International Conference on Communication, Control and Information Sciences (ICCISc). IEEE (2021)
[4] Rahi, Md.M.I., Khan, F.T., Mahtab, M.T., et al.: Detection of skin cancer using deep neural networks. In: IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE). IEEE (2019)
[5] Gulati, S., Bhogal, R.K.: Detection of malignant melanoma using deep learning. In: Advances in Computing and Data Sciences. Springer, pp. 312-325 (2019)
[6] Jojoa Acosta, M.F., Caballero Tovar, L.Y., Garcia-Zapirain, M.B., et al.: Melanoma diagnosis using deep learning techniques on dermatoscopic images. BMC Medical Imaging. 21(1), 6 (2021)
[7] Mukherjee, S., Malu, A., Balamurali, A.R., et al.: TwiSent: A multistage system for analyzing sentiment in Twitter. In: 21st ACM International Conference on Information and Knowledge Management (CIKM '12). ACM (2012)
[8] Sanketh, R.S., Madhu Bala, M., Narendra Reddy, P.V., et al.: Melanoma disease detection using convolutional neural networks. In: 4th International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE (2020)
[9] Attia, M., Hossny, M., Nahavandi, S., et al.: Skin melanoma segmentation using recurrent and convolutional neural networks. In: IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017). IEEE (2017)
[10] Daghrir, J., Tlig, L., Bouchouicha, M., et al.: Melanoma skin cancer detection using deep learning and classical machine learning techniques: A hybrid approach. In: 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP). IEEE (2020)

[11] Al-Rasheed, A., Ksibi, A., Ayadi, M., Alzahrani, A.I., Zakariah, M., Ali Hakami, N.: An ensemble of transfer learning models for the prediction of skin cancers with conditional generative adversarial networks. Diagnostics. 12(12), 3145 (2022)
[12] Natha, P., Sharma, R., Singh, A., et al.: Advancing skin cancer prediction using ensemble models. Computers. 13(7), 157 (2024)
[13] Alotaibi, A., AlSaeed, D.: Skin cancer detection using transfer learning and deep attention mechanisms. Diagnostics. 5(1), 99 (2025)