*Undergraduate Conference on Intelligent Computing and Systems (UCICS 2025)*
*26-27 February, 2025; Varendra University, Rajshahi, Bangladesh*

# A Comparative Analysis of Pre-Trained Convolutional Neural Networks for Melanoma Early Detection

Aniqa Raihana
*Computer Science and Engineering*
*Varendra University*
Rajshahi, Bangladesh
aniqaraihana.99@gmail.com

Md. Shahid Ahammed Shakil
*Computer Science and Engineering*
*Varendra University*
Rajshahi, Bangladesh
shakil.ahammed074gmail.com

Anamika Saha
*Computer Science and Engineering*
*Varendra University*
Rajshahi, Bangladesh
anamikasaha904@gmail.com

Prianka Rani Mukherjee
*Computer Science and Engineering*
*Varendra University*
Rajshahi, Bangladesh
priankamukherjee24@gmail.com

Shamrin Rahman
*Computer Science and Engineering*
*Varendra University*
Rajshahi, Bangladesh
shamrin359@gmail.com

Md. Jamil Chaudhary
*Computer Science and Engineering*
*Varendra University*
Rajshahi, Bangladesh
jamil@vu.edu.bd

***Abstract*—One of the most prevalent and lethal types of cancer in the world is skin cancer, especially melanoma. Improving patient outcomes requires early detection, and automated systems powered by deep learning (DL) have demonstrated significant potential to support medical professionals. Several pre-trained convolutional neural network (CNN) models used for melanoma early detection are compared in this study. They are VGG16, MobileNetV2, EfficientNetB7, and DenseNet121. The HAM10000 dataset, which includes a wide range of skin lesions, was used to assess these models. The aim was to find the best pre-trained model for the classification of melanoma. We compared the accuracy, precision, recall, and F-1 scores of various models to assess their capacity to extract reliable features from the dataset. The CNN (2D) model performed the best, according to the data, in terms of accuracy (87%), precision (86%), recall (85%), and F1-score (88%). By thoroughly contrasting the most widely used pre-trained models and revealing their relative advantages and disadvantages for melanoma diagnosis, this study advances the field. According to the results, these pre-trained models can greatly improve automated melanoma detection, facilitating early diagnosis and offering dermatologists and other healthcare professionals invaluable assistance.**

***Index Terms*—Deep Learning, Machine Learning, Convolutional Neural Network, HAM10000, Skin Lesion.**

## I. Introduction

Artificial neural networks have long been trained using dermatoscopic pictures to identify pigmented skin lesions automatically. The potential of this strategy was demonstrated by Binder et al. in 1994 when they trained a neural network to distinguish between melanocytic nevi and melanomas, the most severe form of skin cancer [1]. One of the most prevalent forms of skin cancer in the world is skin lesions. In the literature, many computerized methods have been introduced to classify skin cancers [2]. For many years, cancer has been the leading cause of death worldwide, accounting for more deaths than any other disease (Baltruschat et al., 2019). In 2018, the world saw a shocking 9.5 million deaths worldwide from cancer. According to predictions, there will be about 1.8 million new instances of cancer in the United States alone in 2020, with an estimated 600,000 people dying from the disease (Barnett et al., 2019) [3]. A 10% reduction in the ozone layer, according to experts, would significantly increase the number of skin cancer cases, resulting in an additional 4,500 melanoma cases and 300,000 non-melanoma cases annually [4]. Early diagnosis greatly enhances melanoma outcomes; however, delayed detection remains a pressing issue due to the worldwide shortage of qualified specialists [5]. To tackle this challenge and advance healthcare for the greater good, CNN technology emerges as a vital solution, offering powerful tools to improve diagnostic accuracy and efficiency [6]. Especially in its early stages, melanoma can be hard to distinguish from benign skin lesions like freckles and moles since they often look alike. This similarity may lead to an incorrect diagnosis if sophisticated diagnostic methods are not used [7]. Small sample sizes and a lack of varied dermatoscopic images plagued early research, mainly focusing on melanomas and melanocytic nevi. Training reliable models that can effectively generalize to different types of lesions is hampered by this restriction [1]. Uncontrollably proliferating and dividing cells are called "cancer"; if treatment is not received, they can swiftly expand and infiltrate neighboring tissues. The greatest chance of turning into a malignant tumor is present in all types of cancer, not just skin cancer [8] [9] [10]. Important organs like the brain, liver, spleen, and lungs can be affected by melanoma, the most deadly kind of skin cancer. The chance of a successful treatment plan is significantly decreased by metastasis, so early detection is crucial [5]. One well-known deep learning model that excels at tackling a variety of challenging tasks over a broad range of fine-grained object categories is the convolutional neural network (CNN) [11] [12]. This study evaluates cutting-edge pre-trained models like VGG16, MobileNetV2, EfficientNetB7, and DenseNet121 extensively to determine the best architecture for skin lesion classification. By carefully evaluating these models on the HAM10000 dataset, the research clarifies the benefits and drawbacks of different pre-trained architectures in the context of medical imaging. By creating a strong and dependable method to categorize skin lesions, especially melanomas, the study focuses on the crucial goal of early melanoma identification. Because it allows for prompt medical action, early and precise melanoma detection is essential for increasing patient survival rates. This work conducts a comprehensive comparison analysis of state-of-threat pre-trained models, including VGG16, MobileNetV2, EfficientNetB7, and DenseNet121, to identify the optimal architecture for skin lesion classification. By carefully evaluating these models on the HAM10000 dataset, the research clarifies the benefits and drawbacks of different pretrained architectures in the context of medical imaging. The work concentrates on the vital objective of early melanoma identification by developing a robust and reliable technique to classify skin lesions, particularly melanomas. Increased patient survival rates depend on early and accurate melanoma detection since it enables timely medical intervention.

## II. RELATED WORK

A deep learning approach designed for melanoma diagnosis using the HAM10000 dataset has achieved remarkable advancements in accuracy and precision. This method integrates a modified ResNet-50 for classification alongside ESRGAN for enhanced image quality [8]. A dual completely convolutional residual network (FCRN) was used for concurrent segmentation and coarse classification in a deep learning system for melanoma diagnosis. A lesion index calculation unit (LICU) was also used to improve classification outcomes. A potential step toward more precise and effective automated melanoma identification in clinical settings was indicated by this system's impressive segmentation, feature extraction, and lesion classification accuracy of 0.753, 0.848, and 0.912 [13].Using data augmentation and models like RegNetY-320, InceptionV3, and AlexNet, another novel framework for skin cancer detection on the unbalanced HAM10000 dataset achieved better results. RegNetY-320 outperformed earlier state-of-the-art techniques with an accuracy of 91%, a ROC value of 0.95, and an F1-score of 88.1% [14]. By using deep convolutional neural network (CNN) characteristics, an automated method for melanoma identification addressed issues such as intra-class variation in skin lesions, inter-class similarity, and limited training data. Using a detailed analysis of the deep characteristics taken from each model, this study assessed how well eight modern CNN architectures classified melanoma. The performance of the technique was evaluated using four benchmark datasets: HAM10000, ISIC 2016, ISIC 2017, and PH2. DenseNet-121 paired with a multi-layer perceptron (MLP) achieved the best accuracy of 98.33%, 80.47%, 81.16%, and 81% on the respective datasets. The method used border localization and normalization techniques. This illustrates how dermatologists can use pre-trained CNNs to help them correctly detect melanoma [15]. In order to address class imbalance and unpredictability in skin cancer categories, EfficientNet models (B0-B7) have been used for multiclass skin cancer classification on the HAM10000 dataset. This includes a preprocessing pipeline with data augmentation, scaling, and hair removal. EfficientNetB4 demonstrated the efficacy of transfer learning with ImageNet weights and fine-tuning by achieving an F1-score of 87% and a Top-1 accuracy of 87.91%. EfficientNetB4 and B5, two models of intermediate complexity, were shown to produce the best results, underscoring the significance of data enhancement and transfer learning in attaining accurate classification [5]. Finally, by doing a thorough comparison examination of pre-trained models, such as VGG16, MobileNetV2, EfficientNetB7, and DenseNet121, for early melanoma diagnosis, this study expands on previous research. Our approach uses the HAM10000 dataset to evaluate multiple designs and select the optimal model for skin lesion classification, unlike previous research that focuses on a particular architecture or specific preprocessing techniques. Our goal is to tackle issues like dataset imbalance and unpredictability in dermatoscopic images by implementing strong data augmentation and transfer learning algorithms. By outlining the benefits and drawbacks of widely used pre-trained models and assisting in the development of a scalable and efficient technique for early melanoma identification, this work closes the gap between research and clinical use.

## III. METHODOLOGIES

### A. Datasets

We used the HAM10000 dataset for our investigation. There are 10,015 high-quality dermatoscopic images of seven different skin lesion types, such as benign keratoses, melanomas, and melanocytic nevi, which are included in the HAM10000 collection. Because of its extensive annotations, which include lesion kind, diagnosis, and location, it is frequently used for automated skin lesion classification. The dataset accommodates a range of model training requirements by offering RGB images at different resolutions. It is an invaluable tool for creating machine learning models for dermatological research and skin cancer diagnosis. It is accessible on Kaggle [16] and needs to be utilized in accordance with data protection and ethics.

### B. Methodology

This study's main goal is to assess how well different pretrained deep-learning models perform in the early diagnosis of melanoma and other forms of skin cancer. To find the best model for this task, a comparison of many pre-trained models was carried out, including VGG16, MobileNetV2, EfficientNetB7, and DenseNet121. The specifics of the approach used in this study are listed below and shown in **Figure** 1.

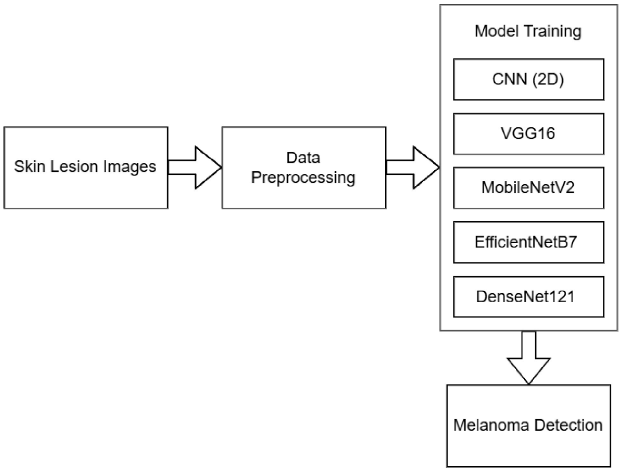


Fig. 1. Melanoma detection framework.

*1) Data Preprocessing:* In this investigation, 10,015 dermatoscopic pictures from seven categories—melanoma, melanoma, nevus, basal cell carcinoma, actinic keratoses, vascular lesions, dermatofibroma, and squamous cell carcinoma—were used from the Skin Cancer MNIST: HAM10000 dataset. These photos are from Kaggle and come with metadata that includes information about the location and diagnosis of the lesion. Several crucial operations were part of the data preparation steps. First, to guarantee consistency for the neural network, every image was shrunk to a standard 150x150 pixel size. The next step was label encoding, which involved using one-hot encoding to encode the categorical values in a binary matrix format once the labels were extracted from the metadata. Image normalization was used to scale the pixel values to a range between 0 and 1 by dividing the pixel values by 255 to standardize the model input. Lastly, using the train test split function from `sci-kit-learn`, the dataset was divided into training and testing sets, with 80% set aside for training and 20% for testing.

*2) Data Augmentation:* To improve the training dataset's diversity and lower the chance of overfitting, data augmentation was used. This method creates altered versions of photos by a variety of modifications, hence artificially expanding the size of the training dataset. Randomly rotating images within a 40-degree range, zooming images within a 30% range, performing horizontal and vertical shifts with a shift range of 20% of the total image width and height, randomly flipping images horizontally to simulate various orientations, and randomly shearing images to simulate angular variations were among the augmentations used in this study. Real-time image augmentation during model training was made possible by the implementation of these changes using Keras' `ImageDataGenerator`. Here are a few stances of augmenting melanoma data. Some examples of our image data augmentation have been shown in **Figure** 2.

*3) Model Selection and Architecture:* Five convolutional neural network (CNN) models that had already been trained were assessed for melanoma detection in this study. The

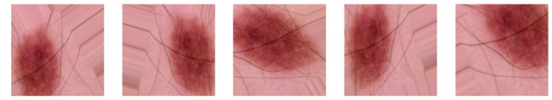

Fig. 2. Some examples of Melanoma data augmentation.

first model is VGG16, a well-known and straightforward CNN architecture that excels at image classification tasks. By substituting a SoftMax function for the last layer, it was adjusted for each of the dataset's seven skin lesion classes. MobileNetV2, the second model, is a thin architecture made for mobile and embedded devices. It is effective in contexts with limited resources because it strikes a compromise between model size and accuracy. A SoftMax layer was added to this model to detect melanoma. The computational efficiency of the fourth model, EfficientNetB7, is acknowledged. It provides good accuracy with minimal computing overhead and scales effectively with growing computer resources. A SoftMax layer was added to this model to categorize skin lesions. Last but not least, DenseNet121 is a lighter, less parameter-heavy variant of DenseNet that enables quicker training and inference. Like the other models, its final layer was adjusted for the seven skin lesion classifications in the dataset, and it still performs competitively despite its smaller size. To make sure these models were appropriate for the melanoma detection job, their final layers were adjusted to fit the dataset's seven types of skin lesions.

*4) Training and Hyperparameter Tuning:* The models were trained using various settings of the hyperparameters. To balance memory consumption and training speed, a batch size of 32 was employed. To avoid overfitting, early stopping was used during the 40 epochs in which the models were trained. Early stopping tracked the validation loss and stopped training after five epochs in which no improvement was seen. For improved gradient descent behavior, the Adam optimizer was employed with a learning rate of 0.001 and beta values of 0.9 and 0.999. Because categorical cross-entropy works well in multiclass classification tasks, it was selected as the loss function. To maximize training, early stopping and learning rate reduction were used. The callback called ReduceLROnPlateau was used in case of the need for degradation of the learning rate.

*5) Model Evaluation:* In order to analyze classification efficacy and error rates, the model's performance was evaluated using accuracy, precision, recall, F1 score, and confusion matrices.

## IV. EXPERIMENTAL RESULT AND ANALYSIS

The goal of this study was to perform a comparative analysis of pre-trained convolutional neural network (CNN) models for the early detection of melanoma. Five pre-trained models—VGG16, MobileNetV2, EfficientNetB7, and DenseNet121—were evaluated on their ability to classify melanoma images. The performance of each model was assessed using standard metrics, including accuracy, precision, recall, and F1-score. The CNN (2D) model achieved the best performance with 87% accuracy, 86% precision, 85% recall, and an 88% F1 score, making it the most reliable model over all. MobileNetV2 and DenseNet121 followed closely with 85% and 82% accuracy, respectively, with MobileNetV2 showing better recall, while DenseNet121 excelled in precision. VGG16 and EfficientNetB7 performed lower, with 80% and 77% accuracy, respectively. However, EfficientNetB7 showed higher precision and recall. A summary of all these findings has been provided in **Table** I. In summary, CNN (2D) is the top performer, while MobileNetV2 and DenseNet121 are suitable alternatives based on specific needs like precision or recall.

TABLE I
PERFORMANCE OF PRETRAINED MODELS ON THE TEST DATASET.

| **Model** | **Accuracy (%)** | **Precision** | **Recall** | **F1-Score** |
|---|---|---|---|---|
| VGG16 | 80 | 79 | 82 | 80 |
| DenseNet121 | 82 | 85 | 80 | 80 |
| MobileNetV2 | 85 | 79 | 80 | 82 |
| EfficientNetB7 | 77 | 82 | 82 | 76 |
| CNN (2D) | 87 | 86 | 85 | 88 |

## V. LIMITATIONS AND FUTURE WORK

Although the CNN (2D) model achieved an accuracy of 87%, there is room for improvement. Future research will focus on regularizing and hyperparameter-tuning models like DenseNet121 and MobileNetV2, as well as exploring ensemble learning strategies to integrate multiple models for enhanced performance. To address overfitting and improve model generalization, data augmentation and synthetic data production will be investigated. Class imbalance, which impacted recall and precision for certain classes, will be addressed using techniques like class weighting or targeted loss. Additionally, testing for latency, scalability, and practical applicability will be necessary for real-world implementation. Despite the models' strong performance in controlled settings, continuous improvement with larger datasets and more efficient methods will be crucial for practical use.

## VI. CONCLUSION

This study compares the performance of pre-trained convolutional neural network (CNN) models—VGG16, DenseNet121, MobileNetV2, EfficientNetB7—and a custom-built 2D CNN in the early diagnosis of melanoma. The custom 2D CNN model demonstrated the highest performance, showcasing the potential of deep learning for enhancing diagnostic accuracy. With melanoma being difficult to distinguish from benign lesions, deep learning models can be invaluable in assisting medical professionals in making precise diagnoses. The research underscores the potential of automated systems to improve early diagnosis and patient outcomes, setting the stage for future advancements in melanoma detection.

## REFERENCES


[1] P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Scientific data*, vol. 5, no. 1, pp. 1-9, 2018, Nature Publishing Group.

[2] Amina Bibi, Muhammad Attique Khan, Muhammad Younus Javed, Usman Tariq, Byeong-Gwon Kang, Yunyoung Nam, Reham R Mostafa, and Rasha H Sakr, "Skin lesion segmentation and classification using conventional and deep learning based framework," *Comput. Mater. Contin.*, vol. 71, no. 2, pp. 2477–2495, 2022.

[3] Girdhar, Nancy and Sinha, Aparna and Gupta, Shivang, "DenseNet-II: An improved deep convolutional neural network for melanoma cancer detection," *Soft Computing*, vol. 27, no. 18, pp. 13285–13304, 2023. Springer.

[4] S. Chaturvedi, K. Gupta, and P. S. Prasad, "Skin lesion analyser: an efficient seven-way multi-class skin cancer classification using MobileNet," in *Advanced Machine Learning Technologies and Applications: Proceedings of AMLTA 2020*, Springer, 2021, pp. 165–176

[5] Ali, Karar, Shaikh, Zaffar Ahmed, Khan, Abdullah Ayub, Laghari, Asif Ali. "Multiclass skin cancer classification using EfficientNets – a first step towards preventing skin cancer." *Neuroscience Informatics*, vol. 2, no. 4, 2022, p. 100034.

[6] K. Young, G. Booth, B. Simpson, R. Dutton, and S. Shrapnel, "Deep neural network or dermatologist?," in *Interpretability of Machine Intelligence in Medical Image Computing and Multimodal Learning for Clinical Decision Support: Second International Workshop, iMIMIC 2019, and 9th International Workshop, ML-CDS 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 17, 2019, Proceedings 9*, Springer, 2019, pp. 48–55.

[7] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115–118. Nature Publishing Group.

[8] Alwakid, G., Gouda, W., Humayun, M., and Sama, N. U., "Melanoma Detection Using Deep Learning-Based Classifications," *Healthcare (Basel)*, vol. 10, no. 12, p. 2481, 2022. doi: 10.3390/healthcare10122481, PMID: 36554004, PMCID: PMC9777935.

[9] Saeed, Jwan, and Zeebaree, Subhi, "Skin lesion classification based on deep convolutional neural networks architectures," *Journal of Applied Science and Technology Trends*, vol. 2, no. 01, pp. 41–51, 2021.

[10] Albahar, Marwan Ali, "Skin lesion classification using convolutional neural network with novel regularizer," *IEEE Access*, vol. 7, pp. 38306–38313, 2019.

[11] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, Nature Publishing Group UK London.

[12] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, *et al.*, "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, pp. 211–252, 2015, Springer.

[13] Li, Yuexiang and Shen, Linlin, *Skin Lesion Analysis towards Melanoma Detection Using Deep Learning Network*, Sensors, vol. 18, no. 2, Article 556, 2018.

[14] Alam, Talha Mahboob, Shaukat, Kamran, Khan, Waseem Ahmad, Hameed, Ibrahim A., Almuqren, Latifah Abd, Raza, Muhammad Ahsan, Aslam, Memoona, and Luo, Suhuai. "An efficient deep learning-based skin cancer classifier for an imbalanced dataset." *Diagnostics*, vol. 12, no. 9, 2022, p. 2115, MDPI.

[15] Himanshu K. Gajera, Deepak Ranjan Nayak, and Mukesh A. Zaveri, "A comprehensive analysis of dermoscopy images for melanoma detection via deep CNN features", *Biomedical Signal Processing and Control*, vol. 79, p. 104186, 2023.

[16] Kevin Mader, *Skin Cancer MNIST: HAM10000*, 2018. Available at: https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000