**Deep Learning-Based Skin Cancer Classification Using ResNet50 and MobileNet Architectures**

Suraj Prasanna  
Dept. of CSE  
Manipal Academy of Higher Education  
Bengaluru, India  
[suraj.mitblr2022@learner.manipal.edu](mailto:suraj.mitblr2022@learner.manipal.edu)

Tanmay Bansal  
Dept. of CSE  
Manipal Academy of Higher Education  
Bengaluru, India  
[tanmay1.mitblr2022@learner.manipal.edu](mailto:tanmay1.mitblr2022@learner.manipal.edu)

Abhinav Dayal  
Dept. of CSE  
Manipal Academy of Higher Education  
Bengaluru, India  
[abhinav2.mitblr2022@learner.manipal.edu](mailto:abhinav2.mitblr2022@learner.manipal.edu)

Anubrat Bora  
Dept. of CSE  
Manipal Academy of Higher Education  
Bengaluru, India  
[anubrat.mitblr2022@learner.manipal.edu](mailto:anubrat.mitblr2022@learner.manipal.edu)

**Abstract**

Skin cancer, one of the most common forms of cancer, can be effectively managed through early detection. In this work, we present a comparative study of two deep convolutional neural network architectures—ResNet50 and MobileNet—for the classification of skin cancer images. The models were trained and evaluated using a labeled dataset of dermoscopic images. Extensive preprocessing, transfer learning, and evaluation techniques were applied. Results demonstrate that both models perform well in classifying cancerous vs non-cancerous lesions, with MobileNet offering a lightweight alternative for deployment on resource-constrained devices. Performance metrics such as accuracy, precision, recall, and ROC curves were used to assess model performance.

# I. Introduction

Skin cancer represents one of the most common and rapidly growing types of cancer worldwide, affecting millions of individuals annually. Its incidence continues to climb due to a combination of environmental factors, such as increased exposure to ultraviolet (UV) radiation, and heightened awareness that leads to more frequent diagnoses. Despite advancements in medical technology, early and accurate detection remains a critical factor in improving patient outcomes and increasing survival rates. Unfortunately, traditional diagnostic methods such as physical examinations, dermoscopic assessments, and confirmatory biopsies are often limited by high costs, time consumption, and dependence on the subjective judgment of medical professionals. These limitations are particularly pronounced in underserved regions where access to dermatologists and specialized diagnostic tools is scarce. In recent years, the field of artificial intelligence has seen transformative growth, especially with the development of deep learning techniques. Among these, convolutional neural networks (CNNs) have emerged as a powerful tool for analyzing and classifying medical images due to their ability to automatically learn and extract complex features from raw data. CNNs have shown tremendous promise in dermatology, offering the potential to assist clinicians in making faster and more reliable diagnoses. This study aims to harness the power of deep learning to address the challenges of skin lesion classification using dermoscopic images. We focus on two state-of-the-art CNN architectures: ResNet50, renowned for its depth and superior performance in image recognition tasks, and MobileNet, a lightweight and efficient model designed for deployment on mobile and embedded systems. Both models are fine-tuned using transfer learning techniques, and their performance is systematically evaluated to assess the balance between classification accuracy and resource efficiency.

II. LITERATURE REVIEW

The application of deep learning to dermatological diagnostics has gained significant traction in recent years. Esteva et al. [1] pioneered the use of deep CNNs for classifying skin lesions at a performance level comparable to dermatologists. Similarly, Codella et al. [2] explored ensemble methods and deep learning for melanoma classification, establishing the importance of combining handcrafted and learned features. In the ISIC challenge, researchers like Tschandl et al. [3] and Kawahara et al. [4] demonstrated the efficacy of CNNs and transfer learning on large dermoscopic image datasets. Yu et al. [5] proposed a very deep residual network approach, showing improvements in melanoma detection accuracy. Mobile-friendly architectures have also gained attention. Howard et al. [6] introduced MobileNet as a lightweight model suitable for mobile and embedded vision applications, making it appealing for deployment in low-resource settings. Tan and Le [7] further advanced this with EfficientNet, balancing model size and performance. ResNet, proposed by He et al. [8], has been extensively used in medical image classification for its depth and residual learning capabilities. Gessert et al. [9] conducted a comprehensive survey and evaluation of CNNs for skin lesion classification, reinforcing the effectiveness of deep transfer learning techniques. Most recently, Brinker et al. [10] highlighted the potential of AI to match and sometimes exceed dermatologist-level classification in real-world clinical settings, advocating for AI-assisted diagnostic tools.

# III. METHODOLOGY

This study employs two prominent convolutional neural network (CNN) architectures: ResNet-50 and MobileNet. ResNet-50 is a 50-layer deep residual network known for its ability to train very deep models efficiently using skip connections, which mitigate the vanishing gradient problem. In contrast, MobileNet is a lightweight architecture optimized for mobile and embedded systems, leveraging depthwise separable convolutions to reduce the number of parameters and computational cost while maintaining competitive accuracy. Both models were initialized with weights pre-trained on the ImageNet dataset to leverage general visual features learned from large-scale image data. To adapt these pre-trained models for binary skin lesion classification, transfer learning was applied. The convolutional base layers of both networks were frozen to preserve the learned feature representations, and a new classification head was added. This custom head included a Global Average Pooling layer, a fully connected dense layer with 128 units and ReLU activation, and a final output layer with a single neuron activated by a sigmoid function, enabling binary output. The models were trained using the Adam optimizer with a learning rate of 0.0001, and the binary cross entropy loss function was used to guide the training process. Model performance was evaluated using accuracy, precision, and recall metrics during both training and validation phases. Additionally, an early stopping mechanism was implemented to monitor validation loss with a patience of three epochs, halting training once no further improvement was observed to prevent overfitting.

IV. DATASET

The dataset used in this study comprises a collection of dermoscopic images of skin lesions, each annotated as either malignant or benign. These images serve as the input for training and evaluating the deep learning models. To ensure compatibility with the input requirements of the selected architectures, all images were resized to a resolution of 224 × 224 pixels. Preprocessing steps were applied uniformly across the dataset, including normalization to scale pixel values between 0 and 1. To enhance model generalization and robustness, data augmentation techniques such as random rotation, horizontal and vertical flipping, and zooming were employed. These augmentations simulate real-world variations in image capture conditions and help reduce the risk of overfitting, especially when training on relatively limited medical imaging datasets.

A collage of several blisters

Description automatically generated

Sample from the skin cancer image dataset

IV. RESULTS

The performance of both models was evaluated using a range of metrics, including accuracy, precision, recall, confusion matrix, and the ROC curve. Both training and validation accuracy suggested that the models generalized well, with minimal overfitting observed. The confusion matrix further confirmed the models' effectiveness in distinguishing between cancerous and non-cancerous skin lesions. Among the two architectures, MobileNet achieved competitive accuracy while consuming significantly fewer computational resources. This makes it an ideal candidate for deployment on resource-constrained or edge devices, such as mobile phones or embedded medical systems.

The ensemble model, created by combining predictions from both ResNet50 and MobileNet, demonstrated a balanced performance with an overall accuracy of **86.68%** and an AUC (Area Under the Curve) of **0.8817**, indicating strong discriminatory ability between malignant and benign cases. The precision score of **82.49%** suggests the model is highly effective at minimizing false positives, making it reliable in clinical contexts where misdiagnosis could lead to unnecessary biopsies. However, the recall score of **48.91%** reveals room for improvement in detecting all positive cases, which is critical in a diagnostic setting. This trade-off between precision and recall highlights the importance of further tuning or potentially incorporating additional diverse models or data to strengthen sensitivity in future ensemble frameworks, as depicted in both the tables below.

**Table 1. Model Comparison**

| **Metric** | **ResNet50** | **MobileNet** |
| --- | --- | --- |
| Final Training Accuracy | 91.5% | 89.3% |
| Final Validation Accuracy | 86.4% | 85.0% |
| Final Training Loss | 0.215 | 0.245 |
| Final Validation Loss | 0.334 | 0.358 |
| Precision | 85.7% | 84.5% |
| Recall | 86.1% | 83.9% |
| F1-Score | 85.9% | 84.2% |
| Inference Time (per image) | ~45 ms | **~18 ms** |
| Model Size & Complexity | Large and computationally intensive | **Lightweight and efficient** |
| Training Stability (Val Curve) | Moderate fluctuations | Relatively stable |
| Deployment Suitability | Suitable for servers or cloud-based systems | **Ideal for mobile and edge devices** |

A graph of a curve

Description automatically generated

Figure 3: ROC Curve

**Table 2. Final Performance Metrics of the Ensemble Model**

| **Metric** | **Value** |
| --- | --- |
| Accuracy | 86.68% |
| Precision | 82.49% |
| Recall | 48.91% |
| AUC | 88.17% |

A blue square with white text

Description automatically generatedThe training and validation performance of both ResNet-50 and MobileNet models is illustrated in Figures 1 ,2 and 3, which plot the confusion matrix, accuracy and loss curves over 10 epochs. As shown in Figure 2, both models demonstrate consistent improvements in training accuracy, with ResNet-50 achieving approximately 91.5% and MobileNet reaching around 89.3% by the final epoch. Validation accuracy trends indicate good generalization for both models, with ResNet-50 slightly outperforming MobileNet at 86.4% versus 85.2%, respectively. Figure 3 presents the loss curves, where ResNet-50’s training loss steadily declines below 0.22, and MobileNet’s decreases to just above 0.25. Validation loss for ResNet-50 also follows a decreasing trend with minor fluctuations, while MobileNet exhibits a slightly more variable loss curve, suggesting less stable generalization. Overall, ResNet-50 demonstrates superior performance in both training and validation phases, while MobileNet offers a commendable balance between efficiency and accuracy, making it particularly suitable for edge deployment scenarios.

A graph of different colored lines

AI-generated content may be incorrect.

**Figure 2. Final Training and Validation Accuracy**

A graph of different colored lines

AI-generated content may be incorrect.

**Figure 3. Final Training and Validation Loss**

IV. Discussion

The results of this study demonstrate that both ResNet-50 and MobileNet are capable of effectively classifying skin lesions as malignant or benign using dermoscopic images. While ResNet-50 achieved slightly higher validation accuracy, MobileNet’s performance was comparably strong despite its significantly smaller size and lower computational requirements. This finding underscores the potential of lightweight architectures for deployment in real-world, resource-constrained environments such as mobile health applications or remote clinics. The use of transfer learning proved to be highly beneficial, allowing the models to converge quickly and achieve strong performance even with a relatively limited medical dataset. Additionally, data augmentation and early stopping helped to mitigate overfitting, resulting in better generalization on unseen data. Although the binary classification task was successfully addressed, the models occasionally misclassified visually ambiguous cases, highlighting the importance of further refinement and possibly incorporating additional patient metadata or multimodal inputs. These outcomes suggest that deep learning holds significant promise for assisting dermatologists and expanding access to early skin cancer screening, particularly in under-resourced areas.

Figure 1. Confusion Matrix

V. Conclusion

This study highlights the potential of deep learning approaches, specifically ResNet-50 and MobileNet, in the automated classification of skin cancer from dermoscopic images. Both models demonstrated strong performance in distinguishing malignant from benign lesions, validating the effectiveness of transfer learning and CNN-based feature extraction in dermatological image analysis. While ResNet-50 achieved slightly higher validation accuracy, MobileNet proved to be a highly efficient alternative, offering a favorable balance between performance and computational cost. Its lightweight architecture makes it particularly suitable for deployment on mobile and edge devices, paving the way for real-time, accessible skin cancer screening tools. Future work could focus on expanding this framework to support multi-class classification of various skin conditions, integrating the models into clinical decision-support systems, and exploring real-time image analysis on mobile platforms. Additionally, further improvements can be achieved through larger and more diverse datasets, ensemble techniques, and collaboration with dermatologists for enhanced model interpretability and clinical validation.

# VI. References

[1] A. Esteva *et al.*, “Dermatologist-level classification of skin cancer with deep neural networks,” *Nature*, vol. 542, no. 7639, pp. 115–118, 2017.

[2] N. Codella *et al.*, “Skin lesion analysis toward melanoma detection: A challenge at the 2017 International Symposium on Biomedical Imaging (ISBI),” *IEEE ISBI*, pp. 168–172, 2018.

[3] P. Tschandl, C. Rosendahl, and H. Kittler, “The HAM10000 dataset: A large collection of multi-sources dermatoscopic images of common pigmented skin lesions,” *Scientific Data*, vol. 5, 2018.

[4] J. Kawahara, A. BenTaieb, and G. Hamarneh, “Deep features to classify skin lesions,” *IEEE ISBI*, pp. 1397–1400, 2016.

[5] L. Yu *et al.*, “Automated melanoma recognition in dermoscopy images via very deep residual networks,” *IEEE TMI*, vol. 36, no. 4, pp. 994–1004, 2017.

[6] A. G. Howard *et al.*, “MobileNets: Efficient convolutional neural networks for mobile vision applications,” *arXiv preprint arXiv:1704.04861*, 2017.

[7] M. Tan and Q. Le, “EfficientNet: Rethinking model scaling for convolutional neural networks,” *ICML*, pp. 6105–6114, 2019.

[8] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” *CVPR*, pp. 770–778, 2016.

[9] N. Gessert *et al.*, “Skin lesion classification using CNNs: A systematic review,” *arXiv preprint arXiv:1809.10388*, 2018.

[10] T. J. Brinker *et al.*, “Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task,” *European Journal of Cancer*, vol. 113, pp. 47–54, 2019.