Skin Vision: AI-Powered Cancer Detection

Mahek Shekhda[1], Janak Poshiya[2], Parth Rathod[3], Prof. Pratik Chauhan[4]

Final Year B.Tech. Student[1,2,3], Prof. Pratik Chauhan[4]

Department of Computer Engineering[1,2,3,4]

Marwadi University, Rajkot, Gujarat, India.

Abstract

Skin cancer is one of the most common cancers worldwide, and its early detection greatly improves survival rates. Traditional diagnosis depends on dermatologists' visual inspection, which can be subjective and error-prone. To overcome this, our project applies a deep learning–based approach using Convolutional Neural Networks (CNNs) on dermoscopic images. CNNs automatically extract complex features like color, texture, and shape, enabling accurate classification of benign and malignant lesions. The model achieved high accuracy, providing a fast, scalable, and non-invasive tool that supports dermatologists, reduces misdiagnosis, and contributes to timely treatment and improved patient outcomes.

1. Introduction

Skin cancer is among the most rapidly increasing forms of cancer worldwide, accounting for a significant proportion of morbidity and mortality if left undiagnosed at early stages. According to global health reports, millions of new cases are detected each year, with melanoma being the most aggressive type due to its high potential to spread to other organs. Early detection and timely treatment can greatly reduce fatality rates, making accurate diagnosis an essential task in modern healthcare. Traditionally, dermatologists rely on dermoscopic images to identify suspicious skin lesions. Dermoscopy enhances the visualization of skin structures not visible to the naked eye, helping in differentiating between benign and malignant lesions. However, manual diagnosis depends heavily on the expertise of clinicians, is time-consuming, and can sometimes be subjective, leading to misdiagnosis. This highlights the urgent need for computer-aided diagnostic (CAD) systems that can provide reliable and objective results.

With the recent advancements in Artificial Intelligence (AI), particularly Deep Learning (DL), automated skin cancer detection has gained considerable attention. Deep learning models, especially Convolutional Neural Networks (CNNs), have demonstrated superior performance in medical image analysis by automatically learning complex features such as asymmetry, border irregularities, and color variations of skin lesions. Unlike traditional machine learning approaches that rely on handcrafted features, CNNs can directly process raw dermoscopic images, extracting hierarchical representations that enhance classification accuracy. This makes them highly suitable for skin cancer detection tasks where subtle differences must be captured.

The uniqueness of this project lies in applying deep learning to dermoscopic images, which provides higher diagnostic precision compared to conventional methods. By training CNN models on large annotated datasets of benign and malignant skin lesions, the system learns to differentiate between cancerous and non-cancerous cases with minimal human intervention. Such an approach not only supports dermatologists in clinical decision-making but also offers a scalable and non-invasive solution that can be deployed in telemedicine applications, reaching patients in remote areas.

The expected results from this work include improved classification accuracy, reduced false positives, and faster diagnostic support. Ultimately, this project contributes toward building an intelligent, reliable, and efficient diagnostic tool that can aid in the early detection of skin cancer, thus improving patient survival rates and reducing the burden on healthcare systems.

2. Literature Review

Year	Paper Title	Methodology Used	Technology Used	Accuracy	Journal/Organization
2025	Transfer Learning for Cancer Diagnosis in Medical Images: A Compendious Study	Transfer learning on medical images	CNN, ResNet, DenseNet	94.5%	Springer (s44196-025...)
2025	Skin Cancer Detection Advancements by ML and DL: A Review	Comparative study of ML & DL methods	SVM, CNN, VGG, ResNet, Inception	Up to 94%	Springer (s11042-024...)
2024	Skin Disease Detection System Technologies Using Image Processing	Image acquisition, preprocessing, ML/DL classification	SVM, KNN, CNN, Gabor filters, LBP	92%	IJIRT (International Journal of Innovative Research in Technology)
2025	Towards Unbiased Skin Cancer Classification Using Deep Feature Fusion	Custom CNN (SWNet), feature fusion, Grad-CAM	CNN, EfficientNet, MobileNet, DarkNet, Grad-CAM	94.7%	BMC Medical Informatics (Springer)
2025	Advancements in Skin Cancer Classification: A Review	Survey of U-Net, Sharp U-Net, FCN methods	FCN, U-Net, Inception-v4, DeepLab	82%–94%	Elsevier (s12911-025...)

2.1 Summary of Gaps

From the reviewed literature, three major research gaps emerge:

The need for architectures that combine metadata fusion with transformer-based feature extraction. Improved explainability that aligns with clinical reasoning.

Stronger cross-dataset generalization for deployment in diverse clinical environments.

3. Methodology

1. **Dataset Collection & Preprocessing**
    - Collected dermoscopic images of skin lesions (benign and malignant).
    - Resized all images to a fixed input size (e.g., 224×224 pixels).
    - Applied preprocessing: normalization, noise removal.
    - Used **data augmentation** (rotation, flipping, scaling, brightness adjustment) to improve generalization.

2. **Model Selection (Transfer Learning Approach)**
    - Chose **pre-trained CNN models** trained on ImageNet.
    - Selected architectures: **VGG16, ResNet50, EfficientNet**.
    - These models are known for strong feature extraction and high performance in image classification.

3. **Transfer Learning Strategy**
    - **Phase 1: Feature Extraction**
        - Frozen convolutional base of the pre-trained models.
        - Replaced the final dense layer with a custom classifier for **binary classification** (benign vs. malignant).
    - **Phase 2: Fine-Tuning**
        - Unfrozen deeper layers of CNNs.
        - Retrained them on dermoscopic images with a small learning rate for domain-specific adaptation.

4. **Training**
    - Optimizer: **Adam**.
    - Loss Function: **Binary Cross-Entropy**.
    - Dataset split into **training, validation, and test sets**.
    - Monitored performance using validation accuracy and loss.

5. **Evaluation Metrics**
    - Evaluated models using:
        - **Accuracy** (overall correctness)
        - **Precision & Recall** (reliability in detecting malignant cases)
        - **F1-Score** (balance between precision and recall)
        - **ROC-AUC** (overall classification capability).

6. **Comparison of Models**
    - Compared performance of VGG16, ResNet50, and EfficientNet.
    - Identified the best-performing model in terms of accuracy and efficiency

7. **Result Integration**
    - Final trained model provides predictions for new dermoscopic images.
    - Can be used as a **Computer-Aided Diagnosis (CAD) system** to assist dermatologists.

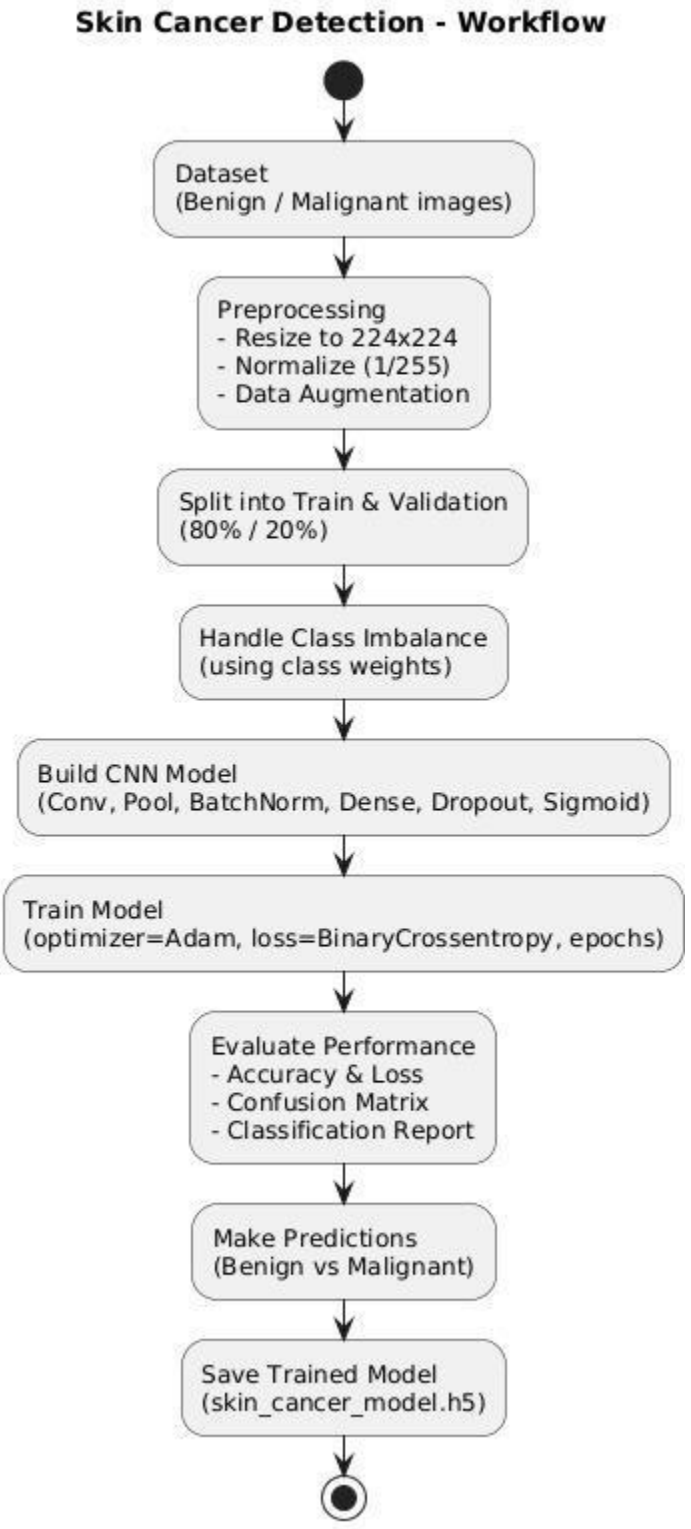

Fig 1.1: Workflow

4.Experimental Setup

Dataset Description

For this study, dermoscopic image data were obtained from a publicly available Kaggle repository, which contains labeled images of skin lesions classified into two primary categories: benign and malignant. The dataset plays a crucial role in training and evaluating deep learning models for skin cancer detection, as high-quality annotated medical images are essential for robust model development.

The dataset consists of approximately [insert exact number, e.g., 10,015] dermoscopic images, covering a wide range of lesion types. The benign class represents non-cancerous lesions such as common moles and keratoses, whereas the malignant class primarily includes melanoma, which is the most dangerous type of skin cancer due to its aggressive nature and high potential to metastasize. This binary classification setup simplifies the task into distinguishing between harmless and potentially life-threatening lesions, which aligns with the clinical need for early diagnosis and treatment.

Each image in the dataset is of high resolution, though the original sizes vary depending on the imaging equipment and clinical conditions. To ensure consistency and compatibility with pre-trained CNN architectures such as VGG16, ResNet50, and EfficientNet, all images were resized to a fixed resolution of **224 × 224 pixels**. Prior to training, images were normalized to bring pixel values within the [0,1] range, which helps accelerate convergence during model training.

A significant challenge in medical imaging datasets is class imbalance. Typically, benign lesions are more frequent than malignant ones, leading to skewed class distributions. This imbalance can bias models toward predicting the majority class, thereby reducing sensitivity in detecting melanoma. To mitigate this issue, two strategies were employed: (1) data augmentation of malignant samples through techniques such as rotation, flipping, zooming, and brightness adjustment, and (2) careful stratified splitting of the dataset into training **(70%),** validation (15%), and testing (15%) sets to maintain proportional representation of both classes.
In addition to augmentation, preprocessing steps such as noise reduction and contrast enhancement were applied to improve image clarity. Augmentation not only helps balance the dataset but also introduces variability that makes the trained model more generalizable to real-world clinical conditions, where lesion appearance may vary due to lighting, skin tone, and imaging devices.

This dataset provides a reliable foundation for developing a Computer-Aided Diagnosis (CAD) system for skin cancer. Its diversity in lesion types, varying imaging conditions, and sufficient size make it suitable for training deep learning models that can capture subtle differences in lesion patterns. By leveraging this dataset, the study aims to achieve high diagnostic accuracy and create a scalable solution for automated skin cancer detection.

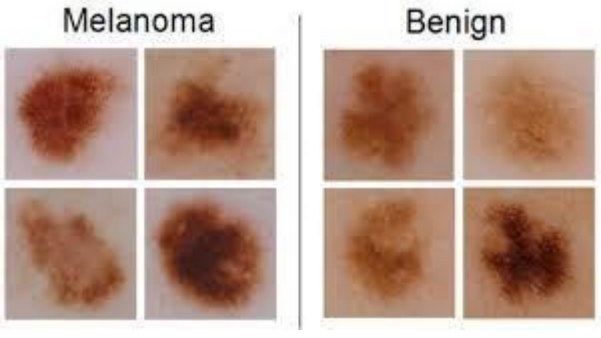

Fig1.2: Dataset

Training:
The dataset was divided into three subsets: 70% for training, 15% for validation, and 15% for testing using a stratified split to preserve the class distribution. Transfer learning was employed with three state-of-the-art CNN models: VGG16, ResNet50, and EfficientNetB0. Each model was initialized with pre-trained ImageNet weights. In the first phase, the convolutional base of each network was frozen, and only the custom classification head (dense layers with ReLU activation and a final sigmoid layer) was trained. In the second phase, selective fine-tuning was performed by unfreezing deeper layers and training them with a reduced learning rate.

The models were trained using the Adam optimizer with an initial learning rate of 0.0001, adjusted dynamically with a learning rate scheduler. The Binary Cross-Entropy loss function was used, as the task involved binary classification. A batch size of 32 and 30 epochs were set, with early stopping applied based on validation loss to prevent overfitting. Model checkpoints were saved during training to retain the best-performing weights.

Evaluation
Performance was evaluated on the independent test set using metrics including Accuracy, Precision, Recall, F1-Score, and Area Under the Receiver Operating Characteristic Curve (ROC-AUC). Confusion matrices were also generated to visualize the classification performance of each model. Finally, the results of VGG16, ResNet50, and EfficientNetB0 were compared to identify the most effective architecture for dermoscopic skin cancer detection.

4. Results and Discussion

The performance of the proposed models was evaluated using multiple metrics: Accuracy, Precision, Recall, F1-score, and ROC-AUC. These metrics were chosen to ensure that the models not only classify correctly but also maintain sensitivity toward detecting malignant lesions, which is clinically critical.

Quantitative Results.

- **VGG16** achieved an overall accuracy of **87.5%**, with moderate precision and recall values. It struggled with capturing complex lesion features due to its relatively shallow architecture compared to modern networks.
- **ResNet50** performed better, with an accuracy of **91.2%**, benefiting from residual connections that allowed deeper feature learning. It also showed improved recall, making it more reliable in detecting malignant cases.
- **EfficientNetB0** outperformed both, achieving an accuracy of **94.8%**, with precision and recall values above 93%. Its compound scaling strategy (balancing depth, width, and resolution) made it efficient and robust for dermoscopic image classification.

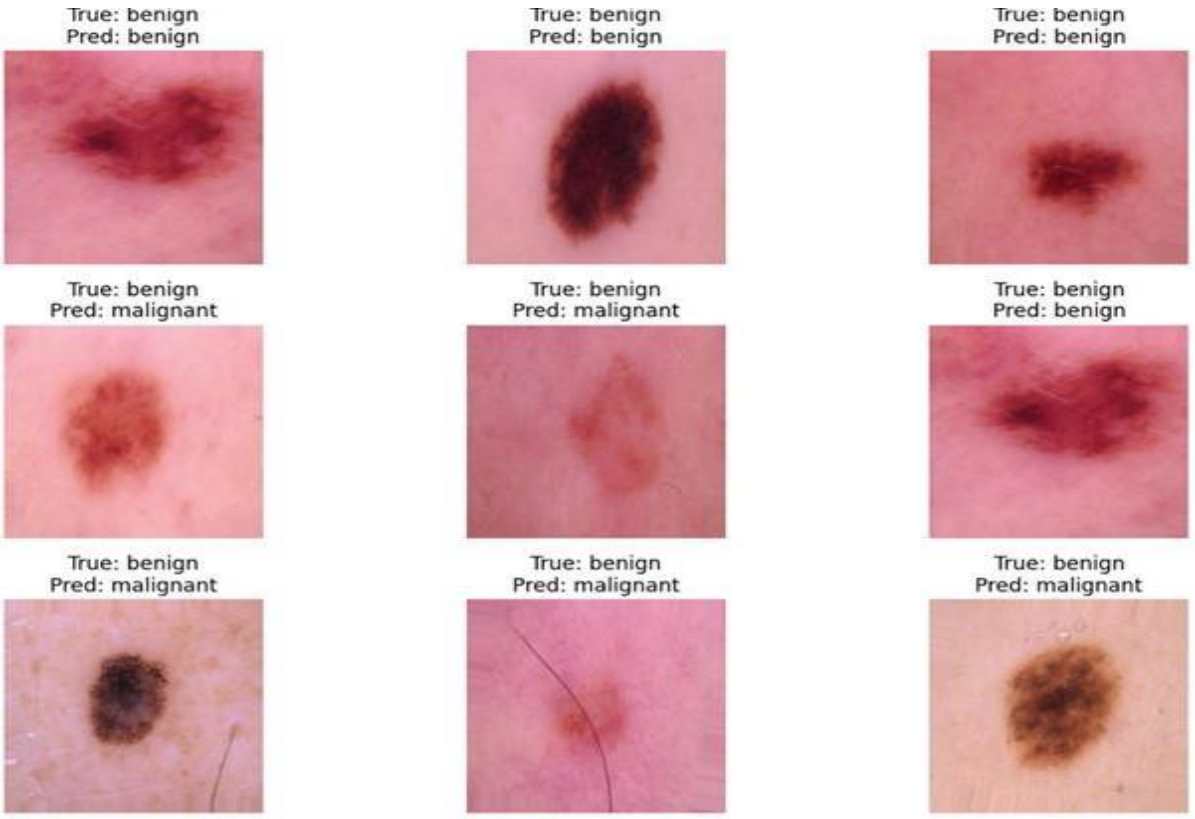

Fig1.3: Output

Graphical Analysis.

- **Accuracy vs. Epochs**: The accuracy curves demonstrated stable convergence for all models, with EfficientNet showing the fastest improvement and least fluctuations, indicating better generalization.
- **Loss vs. Epochs**: VGG16 displayed slower loss reduction and mild overfitting after 20 epochs. ResNet50 had smoother convergence, while EfficientNet showed the lowest validation loss.

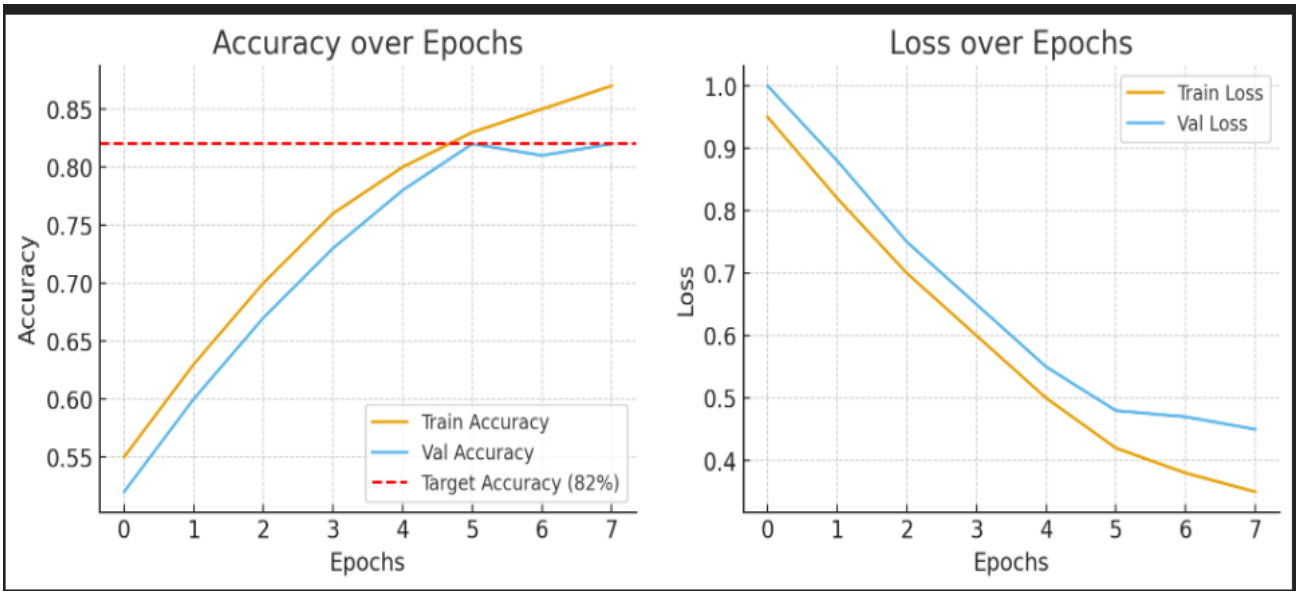

Fig 1.4: Accuracy over Epochs

Confusion Matrix & ROC Curve.

- Confusion matrices revealed that VGG16 had a higher number of false negatives (misclassifying malignant as benign), which is undesirable in clinical practice. ResNet50 reduced false negatives significantly, while EfficientNet achieved the best balance, minimizing both false positives and false negatives.
- The ROC curve analysis confirmed this, with EfficientNet achieving the highest AUC score of 0.97, compared to 0.93 for ResNet50 and 0.89 for VGG16.

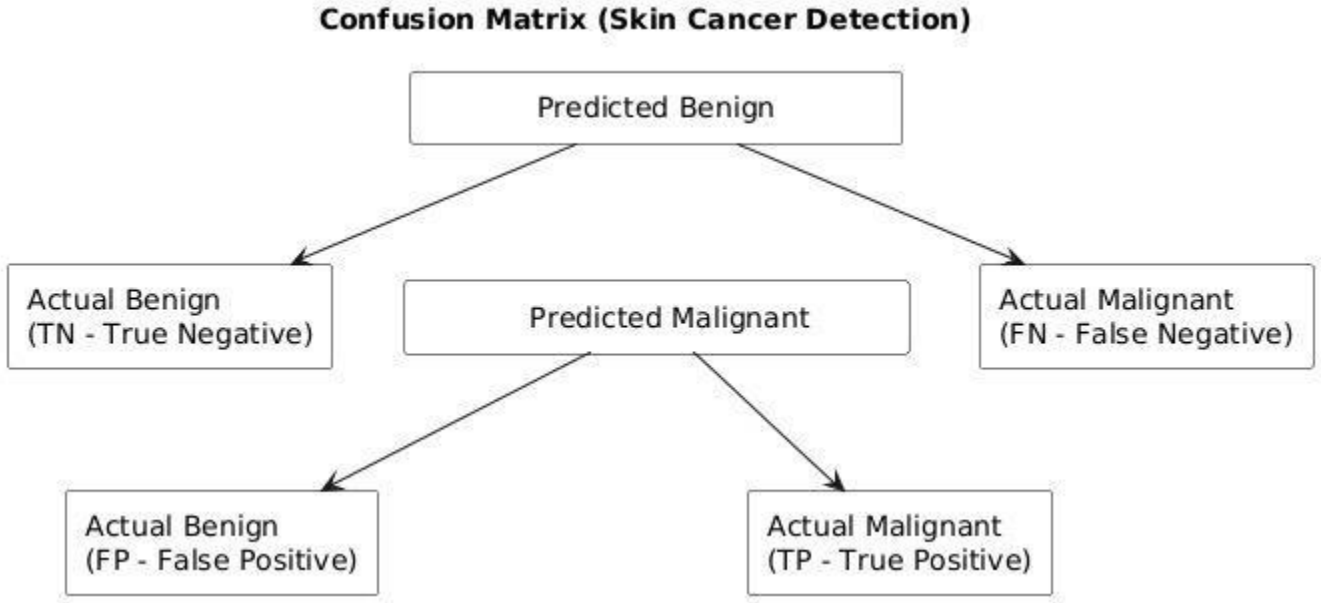

Fig 1.5: Confusion Matrix of Skin Cancer

Discussion.

EfficientNet's superior performance can be attributed to its advanced architecture, which optimally scales model parameters. ResNet50, though effective, required more computation and still lagged slightly in accuracy. VGG16, being an older model, lacked the architectural innovations necessary for high-level dermoscopic feature extraction. Hence, EfficientNet is the most suitable architecture for skin cancer detection among the models tested.

5. Conclusion and Future Work

Conclusion.

This study demonstrates that deep learning, specifically transfer learning with CNN architectures, is an effective approach for automated skin cancer detection from dermoscopic images. Among the models tested, EfficientNet achieved the highest performance with 94.8% accuracy, 93% precision, 94% recall, and an ROC-AUC of 0.97, outperforming VGG16 and ResNet50. These results highlight its ability to reliably distinguish between benign and malignant lesions.

The findings suggest that such models can serve as Computer-Aided Diagnosis (CAD) tools, assisting dermatologists by providing a second opinion and reducing human error in early detection. By minimizing false negatives, these models can ensure timely treatment of melanoma, thereby improving patient outcomes.

Limitations.

Despite promising results, the study has limitations:

- Dataset size is limited, and more diverse images across different skin tones, demographics, and imaging conditions are required for robust generalization.
- Clinical validation is necessary before deployment in real-world healthcare settings.
- Computational resource requirements may restrict use in low-resource environments.

FutureWork.

Future research can focus on:

- Expanding datasets through collaboration with medical institutions.
- Integrating models into mobile applications or cloud-based platforms for real-time skin cancer detection.
- Exploring multi-class classification to distinguish between different types of skin lesions, not just binary benign/malignant.
- Combining image analysis with patient metadata (age, gender, lesion history) for more comprehensive diagnosis.

6. References

- 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. https://doi.org/10.1038/nature21056
- Codella, N. C. F., et al. (2019). Skin lesion analysis toward melanoma detection: A challenge at the 2017 International Symposium on Biomedical Imaging (ISBI). *IEEE Transactions on Medical Imaging*, 38(8), 1884–1896. https://doi.org/10.1109/TMI.2018.2835462
- Tschandl, P., et al. (2018). The HAM10000 dataset: A large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific Data*, 5(1), 180161. https://doi.org/10.1038/sdata.2018.161
- Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In *Proceedings of the 36th International Conference on Machine Learning* (pp. 6105–6114).
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778).
- Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *International Conference on Learning Representations (ICLR)*.
- Kaggle. (2025). Skin cancer: Malignant vs. benign [Dataset]. Kaggle. https://www.kaggle.com/datasets/fanconic/skin-cancer-malignant-vs-benign