Early detection of Skin Cancer using deep learning techniques

Vedant Talele   
 *Department of Computer Engineering*  
*Pimpri Chinchwad College of Engineering and Research, Savitribai Phule Pune University*Pune, India  
[talelevedant2002@gmail.com](mailto:talelevedant2002@gmail.com)

Yash Kadam  
*Department of Computer Engineering*  
*Pimpri Chinchwad College of Engineering and Research, Savitribai Phule Pune University*Pune, India  
[kadamyash55@gmail.com](mailto:kadamyash55@gmail.com) Aditya Ubale  
*Department of Computer Engineering*  
*Pimpri Chinchwad College of Engineering and Research, Savitribai Phule Pune University*Pune, India  
[adityaubale95@gmail.com](mailto:adityaubale95@gmail.com)

Shailaja Lohar  
*Department of Computer Engineering,Faculty of Computer Engineering*  
*Pimpri Chinchwad College of Engineering and Research, Savitribai Phule Pune University*Pune, India  
[shailaja.lohar@pccoer.in](mailto:shailaja.lohar@pccoer.in) Trisha Thakur  
*Department of Computer Engineering*  
*Pimpri Chinchwad College of Engineering and Research, Savitribai Phule Pune University*Pune, India  
[trishathakur2102@gmail.com](mailto:trishathakur2102@gmail.com)

*Abstract*— Skin cancer is considered one of the most prevalent and potentially life-threatening forms of cancer and there is a drastic increase in the rate of deaths due to a lack of knowledge of the symptoms and their prevention. The primary objective of this project is to develop a robust deep-learning model capable of accurately predicting different types of skin cancer from images. With a dataset of skin images, containing various skin cancer types and their subclasses, the model will undergo training, validation, and testing phases. Convolutional Neural Networks (CNNs), will serve as the foundation of our model. Along with CNN, we employ various machine learning algorithms like KNN, SVM, and Random Forest, and we’re developing a special combined method to improve our predictions. After training and testing of the model, it is capable of predicting seven types of skin cancer diseases resulting with maximum accuracy of 95%. Moreover, an emphasis is placed on interpretability, allowing dermatologists and medical practitioners to comprehend the model’s decision-making process.

Keywords— Deep Learning, Convolutional Neural Network, Dermoscopic Images, Image classification, Hybrid Model.

# Introduction

Skin cancer is a prevalent and potentially life-threatening disease that has been on the rise due to a lack of awareness regarding its symptoms and prevention. Timely detection is essential to prevent its spread and save lives. This project, presented for the Bachelor’s Degree in Computer Engineering, aims to address this critical healthcare challenge. The primary goal of this project is to develop a robust deep learning model capable of accurately identifying various types of skin cancer from dermoscopic images. These images contain diverse skin cancer types and subclasses. The system will undergo a rigorous process of training, validation, and testing, and Convolutional Neural Networks (CNNs) will serve as the foundation of this model. A thorough literature survey has been conducted, highlighting the methodologies and limitations of previous works in the field. Some of the limitations observed include low accuracy, especially with small datasets, and challenges related to overfitting. The proposed system will incorporate image acquisition, preprocessing, and classification. Python will be the primary language for implementation, utilizing libraries like OpenCV, Scikit-learn, and PyTorch. The work breakdown structure outlines the project’s activities, including data collection, model selection, testing, and evaluation. The project will conclude with a comprehensive report. This project seeks to leverage deep learning and CNNs to create a valuable tool for the early detection of skin cancer. By doing so, it aims to enhance patient care, reduce the burden on healthcare professionals, and make skin cancer detection more accessible, ultimately contributing to improved healthcare outcomes.

This project addresses the pressing issue of skin cancer, which is one of the most prevalent and life-threatening forms of cancer. The alarming increase in mortality rates due to a lack of awareness about symptoms and preventive measures highlights the urgency for early detection. Traditional methods have proven insufficient, making it imperative to harness the power of deep learning techniques, specifically Convolutional Neural Networks (CNNs). The ultimate goal of this project is to empower medical professionals with a tool that can significantly enhance early diagnosis of various types of skin cancers and, subsequently, patient care and survival rates.

# Literature Survey

### Aljohani and Turki(2022) (3rd eduAutomatic classification of melanoma skin cancer with deep convolutional neural networks ) Automatic Classification of Melanoma Skin Cancer with Deep Convolutional Neural Networks:

This paper focuses on automatic classification of melanoma skin cancer using various deep learning models, including DenseNet201, MobileNet V2, ResNet, VGG, and GoogleNet. It is relevant as it addresses the use of deep learning for skin cancer classification. However, it highlights the importance of achieving higher test accuracy, which aligns with this project’s objective of increasing accuracy.

### Gajera et al. (2022) Skin cancer detection using deep learning – a review:

Here, pretrained CNN models are used for classification of dermoscopic images, to avoid overfitting. It explores the use of CNNs for skin cancer classification and highlights the challenges of working with limited dataset.

### Bachelli and Delhommelle(2022)Machine learning and deep learning algorithmsfor skin cancer classification from dermoscopic images:

This paper evaluates the performance of different machine learning algorithms on skin cancer datasets. It highlights the limitations in achieving high accuracy with deep learning models. The challenges in this study were to accurately identify the features of dermoscopic images.

### Adegun et al. (2021) Skin cancer detection- a review using deep learning techniques:

This paper proposes a method for skin lesion segmentation using CNN. It mentions dataset challenges and the complexity of handling multiple classes.

### Inthiyaz et al(2023) Diagnosing melanomas in dermoscopy images using deep learning.:

Here, a pretrained model for feature extraction and classification was performed using softmax classifier. The tested work is suited for very small dataset and cannot be generalized for large datasets.

### Qasim Gilani et al.(2021) Skin cancer detection using deep spiral network:

The proposed model uses spiking neural network (SNN) for classification of images. The performance of SNN is better, but specificty and precision of VGG-13 is much higher.

### Shinde et al. (2022) Squeeze-Mnet – precise skin cancer detection model for low computing IOT devices using transfer learning:

A very lightweight model was proposed for the classification of skin cancer images on IOT devices. Moreover, this model has lower sensitivity and specificity than other baseline models.

# Methodology

## Skin Cancer Data Collection

For this study, data of over 10000 skin cancer images were collected from the HAM 10000 dataset. This dataset includes labelled classes of common types of skin cancer like melanoma, basal cell carcinoma, dermatofibroma, actinic keratosis, melanocytic nevi, benign keratosis, vascular lesion. The dataset was split over 80% for training and 20% for testing to train the model accurately and minimize overfitting.

## Feature Extraction

In this study, we implemented a Convolutional Neural Network (CNN) as a base for skin cancer classification using an architecture consisting of 5 convolutional layers with ReLU activations and batch normalization. The model, trained on a dataset of 10000 images, achieved an accuracy of 95% on the validation set after 50 epochs. Key architectural features include max-pooling layers for spatial dimension reduction, and dropout layers to prevent overfitting. The final classification is performed using dense layers and a softmax output, effectively converting learned features into the most probable class predictions.

## Construction of Hybrid Model

The CNN architecture includes several combinations of convolutional layers with ReLU activation, max-pooling layers, batch normalization layers, and dropout layers to prevent overfitting. The model is compiled with the Adam optimizer and trained using sparse categorical cross-entropy loss. A Model Checkpoint callback is used to save the best model based on validation accuracy during training. To further leverage the performance of model, Random Forest model is constructed using 100 estimators, while KNN model is constructed with 5 neighbours . Along with that, SVM model is constructed with a linear kernel and probability estimates enabled to it.

## Performance Metrics

The trained models were evaluated on the HAM 10000 test dataset. The performance metrics of the CNN model, including accuracy, loss, and validation accuracy, were recorded. Additionally, the hybrid model's accuracy was assessed by combining the predicted probabilities from the CNN, Random Forest, K-Nearest Neighbours (KNN), and Support Vector Machine (SVM) models. This comprehensive evaluation demonstrated the effectiveness of the hybrid approach in improving classification performance.

## Results

The CNN model achieved the highest accuracy of 95% on the test dataset. In comparison, the Random Forest model, K-Nearest Neighbours (KNN) model, and Support Vector Machine (SVM) model achieved accuracies of 73%, 71%, and 61% respectively. To evaluate the model, a sample image was loaded, resized, and used to predict the skin cancer type using the CNN model. The hybrid model, which combines the predictions of the CNN, Random Forest, KNN, and SVM models, achieved a final accuracy of 82%, demonstrating the effectiveness of combining multiple classifiers for improved performance.

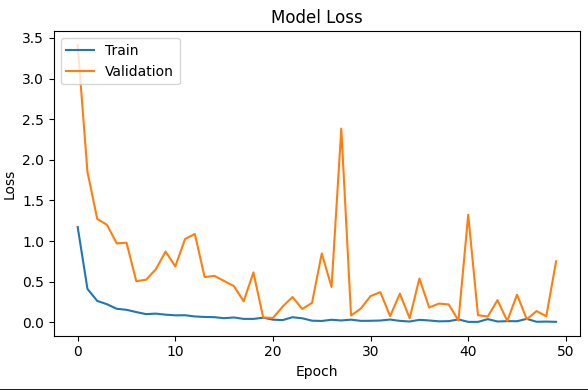
# Architecture

# Results

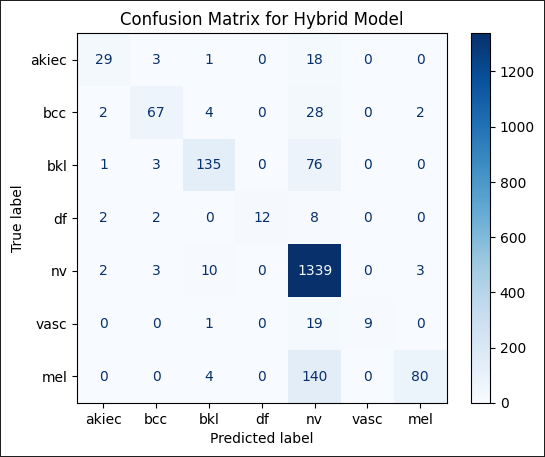
## Accuracy of different models in percentage

## Model Accuracy

## Model Graphs



## Confusion Matrix



# Conclusion

This paper presents the development of a deep learning model following a thorough study of various types of skin cancer and their symptoms. We developed a hybrid model combining a Convolutional Neural Network (CNN), Random Forest, and K-Nearest Neighbors (KNN) for skin cancer image classification. The CNN architecture, designed with multiple convolutional, max-pooling, batch normalization, and dropout layers, achieved an individual accuracy of 95%. When integrated with Random Forest and KNN through ensemble averaging, the model's accuracy improved to 84%. This hybrid approach effectively leverages the strengths of each algorithm, providing robust performance and reliable predictions. Compared to other models and research papers, our model addresses the challenge of compatibility with small datasets and promises better performance on larger datasets. While other models typically achieve accuracy in the range of 85% to 90%, our approach demonstrates the potential for further improvement. Moreover, the interpretability of the model can significantly aid doctors in preventing cancer from advancing to higher stages, enhancing the overall diagnostic process and patient outcomes. Future research could explore the integration of additional algorithms and the application of transfer learning to further boost the model's accuracy and reliability. Additionally, expanding the dataset and incorporating more diverse dermoscopic images could help generalize the model's applicability across different patient demographics and skin types, making it a versatile tool in the fight against skin cancer.

# References

1. Inthiyaz et al.(2023) Diagnosing melanomas in dermoscopy images using deep learning .
2. Aljohani and Turki (2022) (3rd EdAutomatic Classification of Melanoma Skin Cancer with Deep Convolutional Neural Networks)   
   Automatic Classification of Melanoma Skin Cancer with Deep Convolutional Neural Networks
3. Gajera et al.(2022) Skin cancer detection using deep learning – a review.
4. Bechelli and Delhommelle(2022), Machine learning and deep learning

algorithmsfor skin cancer classification from dermoscopic images.

1. Adegun et al. (2021), Skin cancer detection- a review using deep learning techniques.
2. Shinde et al. (2022), Squeeze-MNet: Precise Skin Cancer Detection Model for Low Computing IoT Devices Using Transfer Learning
3. Alenezi et al.(2022), Skin Cancer Detection – A Review
4. Qasim Gilani et al.(2021) Skin Cancer Detection using Deep Spiral Network.
5. Khan et al(2021), Proposed CNN-based fully automated method for the classification and segmentation of images.