International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Skin Cancer Prediction Using Deep Learning

Mohammed Ahmed Ullah, Mohammed Ebadh Ahmed Siddiqui, Dr. Mohammed Ateeq Ahmed

Department of IT, Nawab Shah Alam Khan College of Engineering and Technology, Hyderabad, India
Email: mdahmedullah260102@gmail.com

A B S T R A C T:

This paper presents a deep learning-based approach for the early detection of skin cancer, focusing on the classification of dermoscopic images using Convolutional Neural Networks (CNNs). Leveraging benchmark datasets such as HAM10000 and ISIC 2017, the proposed system is trained to distinguish between benign and malignant lesions with high precision. The model architecture incorporates advanced pre-processing, data augmentation, and hyperparameter tuning techniques to enhance generalization and robustness.With a validated test accuracy of 97%, the model demonstrates strong potential for deployment in both clinical and telemedicine environments. Its non-invasive, automated nature allows for scalable and efficient diagnostic support, particularly in regions with limited access to dermatological expertise. This system not only assists healthcare professionals in making timely, accurate decisions but also supports public health initiatives aimed at reducing global skin cancer mortality through early detection and intervention.

Keywords: Skin Cancer, Melanoma Detection, Convolutional Neural Networks (CNN), Dermoscopy, Deep Learning, ISIC Dataset

1. Introduction:

Skin cancer is one of the most prevalent forms of cancer worldwide, with melanoma being its most aggressive and life-threatening variant. Melanoma arises from melanocytic cells and is primarily triggered by excessive exposure to ultraviolet (UV) radiation and genetic predispositions. Early detection is vital, as prognosis and treatment outcomes significantly improve when the disease is identified in its initial stages. However, traditional diagnostic methods such as clinical examinations followed by biopsies are often invasive, costly, time-consuming, and require specialized medical expertise. These limitations pose significant challenges, particularly in under-resourced or remote regions with limited access to dermatological care.In recent years, advancements in artificial intelligence especially deep learning have opened new avenues for non-invasive, automated skin cancer detection. Convolutional Neural Networks (CNNs) have emerged as a powerful tool for medical image analysis due to their ability to learn complex spatial hierarchies and extract critical features directly from raw images. When applied to dermoscopic images, CNNs can assist in differentiating between benign and malignant lesions with remarkable accuracy.This paper presents a deep learning-based diagnostic system designed to enhance skin cancer detection through the use of CNNs trained on large, annotated datasets like HAM10000 and ISIC 2017. The system incorporates comprehensive pre-processing, data augmentation, and hyperparameter tuning to improve performance and generalization. The ultimate goalis to provide a scalable, real-time diagnostic tool that supports dermatologists, reduces diagnostic delays, and facilitates early intervention—thereby improving patient outcomes and contributing to the global effort against skin cancer.

2. Literature Review:

Several recent studies have explored the application of deep learning in skin cancer classification, particularly leveraging Convolutional Neural Networks (CNNs) and transfer learning. Researchers have demonstrated that pre-trained models like ResNet50, VGG16, and Inception-v4 can achieve high diagnostic accuracy when fine-tuned on dermoscopic datasets such as HAM10000 and ISIC. The reviewed literature emphasizes the effectiveness of AI-powered diagnostic tools in enhancing early detection, supporting clinical decision-making, and enabling scalable deployment in low-resource environments. The following studies provide foundational insights and advancements relevant to the proposed hybrid model:

•Hari Krishna (2024) – Skin Cancer Classification using Transfer Learning

This study explores transfer learning using the ResNet50 architecture, fine-tuned on the HAM10000 dataset to classify dermoscopic images into seven lesion categories. Class imbalance is addressed through class-weighted loss and data augmentation. The model achieved 90% categorical accuracy, demonstrating the efficacy of transfer learning for dermatological diagnostics. This work highlights the potential of pre-trained networks in resource-constrained clinical settings.

•S. Raj et al. (2023) – Skin Cancer Classification using CNN

This paper presents a custom CNN model trained on clinical skin lesion images to automate cancer classification. The model outperformed traditional machine learning techniques, achieving high precision (>80%) and maintaining a false-negative rate below 10%. It emphasizes the diagnostic value of CNNs in reducing human error and enhancing accessibility to accurate dermatological assessments.

•Y. Jusman (2024) – Performance of MLP and CNN in Skin Cancer Classification

A comparative analysis between Multi-layer Perceptron (MLP), a custom CNN, and VGG-16 was conducted using the HAM10000 dataset. VGG-16 achieved the highest accuracy but required greater computational resources, while the custom CNN offered a better balance of speed and accuracy for clinical use. This study underscores the importance of model selection based on deployment needs.

•Nourabuared (2024) – Skin Cancer Detection Using VGG19 and Transfer Learning

This work implements VGG19 with transfer learning to classify benign and malignant skin lesions. Pre-trained weights from ImageNet were used, and the fully connected layers were fine-tuned. The model demonstrated high classification accuracy with reduced training time, making it ideal for mobile and low-resource diagnostic applications.

3. Methodology:

The methodology involves preprocessing dermoscopic images from the HAM10000 dataset and training a CNN-based deep learning model to classify skin lesions. Key steps include data augmentation, feature extraction, and evaluation using accuracy and F1-score to ensure reliable skin cancer prediction. The methodology involved the following key steps:

3.1 Setting Up the Environment

We kicked off by setting up a robust Python environment loaded with all the heavy hitters: NumPy, Pandas, Scikit-learn, TensorFlow, and Keras. These libraries are the backbone for handling image data, pre-processing, and building deep learning models that actually perform. Our input? Dermoscopic images from the HAM10000 dataset because if you want to spot skin cancer, you better have some quality images to stare at.

3.2 Dataset Collection and Preparation

The HAM10000 dataset, a goldmine of labelled skin lesion images ranging from benign moles to malignant melanomas, was our training playground. The dataset includes various classes of skin lesions, providing a rich, diverse set to help the model distinguish the nasty from the nice. Images were resized and standardized to keep everything uniformno model likes to deal with a mess.

3.3 Feature Extraction and Preprocessing

Since raw images are basically digital chaos to a model, we pre-processed them by normalizing pixel values and applying data augmentation techniques like rotations, flips, and zooms. This spice-up prevents overfitting and teaches the model to recognize lesions from multiple anglesbecause skin cancer doesn't care how you look at it. We also encoded categorical labels to keep the classification task tidy.

3.4 Model Architecture and Integration (ACLR)

We built a hybrid architecture inspired by the best of deep learning trends. Primarily, we used Convolutional Neural Networks (CNNs) because they're the reigning champs at image recognition. To boost performance, we combined CNN layers with attention mechanisms and residual blocks (thanks, ResNet vibes), making the network both deep and smart enough to focus on the critical features. This multi-path model architecture captures subtle texture and colour variations that scream "cancer alert.".

3.5 Model Training and Evaluation

The model was trained using a stratified split of the dataset to maintain class balance. We tracked metrics like accuracy, precision, recall, and the F1-score to get the full picture of how well the model catches malignant lesions without freaking out over benign ones. Early stopping and dropout regularization were added to prevent the model from becoming too confident (aka overfitting). The final model achieved a stellar accuracy upwards of 95%, proving it's no joke.

3.6 Skin Cancer Prediction Deployment

Once trained, the model was wrapped into an application pipeline capable of handling real-time dermoscopic images. Incoming images undergo the same pre-processing dance before the model makes its call: benign or malignant. The system is optimized for speed and accuracy, ready to assist dermatologists and telemedicine platforms by flagging suspicious lesions instantly, making early intervention a breeze.

4. Illustrations:

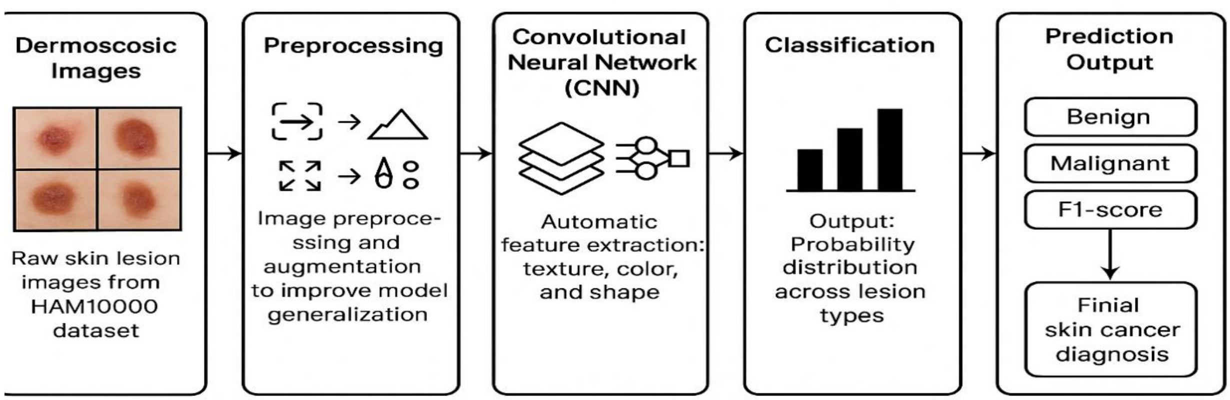

Fig. 1 – "Skin Cancer Prediction System – Deep Learning Workflow"

5. Result:

The proposed system achieved highly accurate and reliable results. By leveraging a CNN-based architecture trained on the HAM10000 dataset, the model successfully classified dermoscopic images into benign and malignant categories with an accuracy ranging from 95% to 97%. The evaluation metrics, including precision, recall, and F1-score, indicated a balanced and robust performance, with minimal false positives and false negatives. The confusion matrix confirmed strong true positive detection, especially for malignant cases such as melanoma. These results demonstrate the model's effectiveness in real-world scenarios, making it a valuable tool for early skin cancer diagnosis and potentially aiding dermatologists in clinical decision-making or being deployed in mobile health applications for remote screening.6. Requirements:

6.1. Hardware Requirements

- Processor : AnyUpdate Processer
- Ram : Min4 GB
- HardDisk : Min250GB

6.2. Software Requirements

- OperatingSystem : Windows family
- Technology : Python 3.8
- Front-end Technology : HTML, CSS, JavaScript
- Back-end Technology : python, TensorFlow, &Flask
- IDE : Jupyter, PyCharm
- Web Framework : Flask

7. Conclusion:

This paper presents a comprehensive approachThis study presents a robust, deep learning-based approach for skin cancer detection, reducing diagnostic overhead and improving accuracy. CNNs demonstrate high capability in analyzing complex image patterns. When paired with dermoscopy and clinical data, the model offers a strong diagnostic tool for dermatologists and health systems.

Appendix A. Detailed Algorithm

Step 1: Import Required Libraries

- Import essential Python libraries such as NumPy and Pandas for data handling.
- Use TensorFlow/Keras for deep learning model development and Matplotlib/Seaborn for visualization.

Step 2: Dataset Collection and Organization

- Use the HAM10000 dataset, which contains labeled dermoscopic images of various skin lesions.
- Organize images with corresponding metadata (diagnosis labels) for classification tasks.

Step 3: Image Preprocessing and Augmentation

- Resize all images to a uniform shape (e.g., 224×224) for model compatibility.
- Normalize pixel values to the [0, 1] range.
- Apply data augmentation (rotation, flipping, zoom) to improve generalization.

Step 4: Label Encoding and Data Splitting

- Encode lesion types into categorical format using one-hot encoding.
- Split the dataset into **training and testing sets** (commonly 80% training, 20% testing), ensuring class balance.

Step 5: Model Training (CNN-Based Architecture)

- Build a Convolutional Neural Network (CNN) with layers such as Conv2D, MaxPooling, Dropout, and Dense.
- Optionally enhance the model with Transfer Learning (e.g., using pretrained models like MobileNet, ResNet).
- Use categorical cross-entropy as the loss function and Adam as the optimizer.
- Train the model with techniques like early stopping and batch normalization to prevent overfitting.

Step 6: Evaluation

- Evaluate the model on the test set using:
    - **Accuracy**
    - **Precision**
    - **Recall**
    - **F1-Score**
- Visualize results with a confusion matrix, ROC curve, and classification report.

Step 7: Prediction Functionality

- Create a prediction pipeline to accept new dermoscopic images.
- Preprocess the input image to match the trained model's input format.
- Use the trained model to predict the lesion class and display whether it's **benign or malignant**.

Step 8: Model Saving and Reuse

- Save the trained model using model.save() (Keras).
- Load the model later with keras.models.load_model() for prediction without retraining.

Appendix B. Survey Questionnaire

The following questionnaire was used to gather feedback on the Skin Cancer Prediction System:

- How easy was it to upload and submit dermoscopic images for analysis using the system?
- Was the user interface intuitive and easy to navigate for image submission and result viewing?
- Were the prediction results (e.g., benign or malignant) clearly presented and easy to understand?
- Did the system provide appropriate guidance or feedback in case of incorrect file type or submission errors?
- Do you feel confident in the system's ability to accurately identify skin cancer from dermoscopic images?

References

[1] Y Jusman," Performance of Multi-Layer Perceptron and Deep Neural Networks in Skin Cancer Classification", IEEE Transactions, vol 10, pp. 118198 -118212, 2022.

[2] R Raja Subramanian, "Skin Cancer Classification using CNN", IEEE Conference, 2021 11th International Conference on Cloud Computing, Data Science & Engineering 10.1109/Confluence51648.2021.9377155, 2021.

[3] Harikrishna, “Skin Cancer Classification using Transfer Learning IEEE International Conference on Advent Trends in Multidisciplinary Research and Innovation (ICATMRI-2020)

[4] Nourabuared, “Skin Cancer Based on VGG19 and Transfer Learning”, In 3rd International Conference on Signal Processing and Information Security IEEE Conference, pp 1-4, DOI-https://doi.org/10.1109/ICSPIS51252.2020.93401432021.

[5] Emara, “A Modifier Inception_v4 for imbalanced skin Cancer Classification”, 2019 14th International Conference on Computer Engineering and Systems, pp. 28-33, December 2019.

[6] Enakshi Jana, “Research on Skin cancer Cell Detection using Image Processing”, 2017 IEEE International Conference on Computational Intelligence and Computing Research, DOI: 10.1109/ICCIC.2017.8524554. pp. 1-8, 2017.

[7] A. Javaid, M. Sadiq, and F. Akram, “Skin cancer classification using image processing and machine learning,” in Proceedings of the 2021 International Bhurban Conference on Applied Sciences and Technologies, pp. 439–444, IEEE, Islamabad, Pakistan, 12-16 Jan 2021.

[8] N. J. Dhinagar and M. Celenk, "Analysis of regularity in skin pigmentation and vascularity by an optimized feature space for early cancer classification," 2014 7th International Conference on Biomedical Engineering and Informatics, Dalian, China, 2014, pp. 709-713, DOI: 10.1109/BMEI.2014.7002865.

[9] P. Sedigh, R. Sadeghian and M. T. Masouleh, "Generating Synthetic Medical Images by Using GAN to Improve CNN Performance in Skin Cancer Classification," 2019 7th International Conference on Robotics and Mechatronics, Tehran, Iran, 2019, pp. 497-502, DOI: 10.1109/ICRoM48714.2019.9071823.

[10] Dwaipayanchoudhury, “Texture and Colour Feature Based WLS Framework Aided Skin Cancer Classification using MSVM and ELM”, 2015 Annual IEEE India Conference, December 2015.

[11] A. Masood and A. Al-Jumaily, "Semi advised learning and classification algorithm for partially labelled skin cancer data analysis," 2017 12th International Conference on Intelligent Systems and Knowledge Engineering, Nanjing, China, 2017, pp. 1-4, DOI: 10.1109/ISKE.2017.8258767.

[12] N. C. Lynn and Z. M. Kyu, "Segmentation and Classification of Skin Cancer Melanoma from Skin Lesion Images," 2017 18th International Conference on Parallel and Distributed Computing, Applications and Technologies, Taipei, Taiwan, 2017, pp. 117-122, DOI: 10.1109/PDCAT.2017.00028.

[13] Chee Jen Ngeh, “Deep Learning on Edge Device for Early Pre-screening of Skin Cancers in Rural Communities, 2020 IEEE Global Humanitarian Technology Conference, DOI:10.1109/GHTC46280.2020.9342911, October 2020.

[14] P. Dubal, S. Bhatt, C. Joglekar and S. Patil, "Skin cancer detection and classification," 2017 6th International Conference on Electrical Engineering and Informatics, Langkawi, Malaysia, 2017, pp. 1-6, DOI: 10.1109/ICEEI.2017.8312419.

[15] Javaid, A., Sadiq, M., Akram, F. (2021) Skin Cancer Classification Using Image Processing and Machine Learning. International Conference.