Rohini A Zambare, 2025, 13:3
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology

An Open Access Journal

Skin Cancer Prediction Using Machine Learning

Rohini A Zambare, N.S.Kulkarni
Siddhant college of engineering Pune

Abstract- Skin cancer is one of the most commonly diagnosed cancers worldwide, and early detection significantly improves treatment outcomes. Machine learning (ML) has emerged as a powerful tool in medical diagnostics, offering accurate, efficient, and scalable solutions for skin cancer prediction. This paper presents a comprehensive approach to classifying skin lesions using ML models such as Convolutional Neural Networks (CNNs), Support Vector Machines (SVMs), and Random Forests. Using datasets like HAM10000 and ISIC, we analyze performance metrics including accuracy, precision, recall, and F1-score. The experimental results show that CNN-based models outperform traditional ML algorithms in detecting melanoma and other skin cancers. This study demonstrates the potential of AI-assisted dermatological diagnosis, thereby contributing to improved clinical workflows and patient outcomes.

Keywords- Skin cancer, melanoma, machine learning, convolutional neural networks, dermatology, classification, HAM10000, medical imaging.

I. INTRODUCTION

Skin cancer, particularly melanoma, poses a significant public health threat due to its high mortality rate if not detected early. Traditional diagnosis relies on dermatoscopic evaluation, which is subjective and time-consuming. With the rise of artificial intelligence (AI) and machine learning, there is increasing interest in automating the detection and classification of skin lesions. Machine learning enables data-driven predictions, reducing diagnostic errors and aiding early detection. This paper focuses on ML models applied to dermoscopic images to predict skin cancer with high accuracy.

II. RELATED WORK

Numerous studies have employed ML techniques for skin cancer classification:

- Codella et al. proposed deep learning models for melanoma detection using ISIC datasets.
- Tschandl et al. presented a multi-class skin disease classification benchmark.
- Kawahara et al. used CNNs for feature extraction from dermoscopy images, enhancing classification performance.

III. DATASETS

We use the HAM10000 and ISIC 2018 datasets:

- **HAM10000:** Contains 10,000+ dermatoscopic images across 7 classes of skin lesions.
- **ISIC:** A well-known dataset from the International Skin Imaging Collaboration.

All images are resized to 224x224 pixels for model compatibility.

IV. METHODOLOGY

Preprocessing

- Image normalization
- Data augmentation (flipping, rotation, zoom)
- Noise reduction using Gaussian filtering

Machine Learning Models

- Support Vector Machine (SVM) – effective for binary classification.
- Random Forest (RF) – ensemble-based method for multi-class prediction.
- Convolutional Neural Network (CNN) – deep learning model that extracts spatial features.

CNN Architecture

We used a custom 5-layer CNN:

- Conv-ReLU-MaxPool
- Batch Normalization
- Dropout (0.3)
- Fully connected layers with Softmax activation

V. EXPERIMENTAL RESULTS

Model	Accuracy	Precision	Recall	F1-score
SVM	84.3%	82.9%	83.1%	82.8%
Random Forest	86.7%	85.5%	85.9%	85.7%
CNN	92.1%	91.3%	91.8%	91.6%

VI. DISCUSSION

CNNs significantly outperform traditional ML models in skin lesion classification. The major challenges include:

- Imbalanced data (fewer malignant samples)
- Visual similarity between lesion types
- Requirement of large labeled datasets

Addressing these with transfer learning, data augmentation, and ensemble models can further boost performance.

VII. CONCLUSION

This research demonstrates that machine learning, especially CNNs, holds promise for automated skin cancer detection. With the increasing availability of dermoscopic image data and advancements in computational power, such ML systems can assist dermatologists in early and accurate diagnosis. Future work will explore hybrid models combining CNNs with transformers and integration into clinical mobile applications.

References

1. Codella, N. C., et al. 'Skin lesion analysis toward melanoma detection: A challenge at the 2017 ISIC.' IEEE CVPR Workshops, 2018.
2. Tschandl, P., et al. 'The HAM10000 dataset: A large collection of multi-source dermatoscopic images of common pigmented skin lesions.' Scientific Data, 5(1), 2018.
3. Kawahara, J., et al. 'Seven-point checklist and skin lesion classification using multi-task multi-modal neural nets.' IEEE Journal of Biomedical and Health Informatics, 2019.
4. Esteva, A., et al. 'Dermatologist-level classification of skin cancer with deep neural networks.' Nature, 542, 2017.
5. Brinker, T. J., et al. 'Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task.' European Journal of Cancer, 2019.
6. Liu, Y., et al. 'A deep learning system for differential diagnosis of skin diseases.' Nature Medicine, 26(6), 2020.
7. Mahbod, A., et al. 'Fusing fine-tuned deep features for skin lesion classification.' Computers in Biology and Medicine, 2020.
8. Haenssle, H. A., et al. 'Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists.' Annals of Oncology, 2018