International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 01 | Jan - 2025 SJIF Rating: 8.448 ISSN: 2582-3930

Skin Lesion Detection and Classification using Convolutional Neural Network

Sushmitha D
Dept.of.CSE, PESITM
Shimogga, India
sushmithadevraj2003@gmail.com

Thejaswini K R
Dept.of.CSE, PESITM
Shimogga, India
thejaswinikr@gmail.com

Varshini M G
Dept.of.CSE,PESITM
Shimogga, India
varshinimg2003@gmail.com

Varshini M V
Dept.of.CSE, PESITM
Shimogga, India
varshini1779@gmail.com

Gagana nayaka
Dept.of.CSE, PESITM
Shimogga, India
gaganagr@pestrust.edu.in

***Abstract* – Skin cancer, particularly melanoma, poses a major public health challenge due to its high mortality rate when not detected at an early stage. Computer-aided diagnosis (CAD) systems have emerged as valuable tools to assist dermatologists in accurately diagnosing skin lesions. This study introduces a system based on convolutional neural networks (CNNs) for detecting and classifying skin lesions. The proposed framework incorporates preprocessing techniques, such as image enhancement and noise reduction, to improve the clarity and visibility of lesions. This is followed by deep learning-based feature extraction and classification. The CNN model is trained using publicly available datasets like HAM10000, which provides a diverse collection of dermoscopic images representing various types of lesions. The system demonstrates high accuracy in differentiating malignant from benign lesions, underscoring the potential of deep learning in medical image analysis. By facilitating early detection, this approach seeks to serve as a reliable and efficient tool, improving treatment outcomes and enhancing patient survival rates..**

Keywords* – Skin lesion detection, convolutional neural network, Melanoma, computer-aided diagnosis, Deep learning.

I. Introduction

Skin cancer is one of the most prevalent forms of cancer globally, with melanoma being the most lethal type. The World Health Organization (WHO) reports a continuous rise in the global incidence of skin cancer, highlight the necessary of early detection and diagnosis for effective treatment and improved survival rates. While melanoma is highly treatable when diagnosed early, delayed identification often leads to poor outcomes and increased mortality. Traditional diagnostic methods rely predominantly on the expertise of dermatologists, which can be subjective and susceptible to errors, especially in settings with limited medical resources.

Advancements in artificial intelligence (AI) and deep learning have introduced promising solutions to these challenges, offering automated, accurate, and scalable approaches for medical diagnosis. Among these, Convolutional Neural Networks (CNNs) have emerged as a leading technology for tasks such as image detection, segmentation, and classification. CNNs are particularly effective for analyzing dermoscopic images because they can automatically learn and extract hierarchical features directly from raw image data, removing the need for manual feature engineering.

The major goal of this project is to build a system with CNN architecture to detect and classify skin lesions. The proposed framework categorizes dermoscopic images into distinct classes, such as malignant (e.g., melanoma) and benign (e.g., nevus) lesions, assisting clinicians in making timely and accurate diagnoses. To improve the model's performance, preprocessing techniques like image augmentation, resizing, and noise reduction are incorporated. The system is trained and evaluated using the HAM10000 dataset, a widely recognized benchmark in skin cancer research.

The application of CNNs in dermatology has the potential to significantly reduce diagnostic errors, enhance access to high-quality healthcare, and improve patient outcomes. This study emphasizes the development and methodology of CNN-based systems for skin lesion detection and classification, highlighting their role as a supportive tool for dermatologists in clinical decision-making.

II. Related Work

The application of deep learning, particularly convolutional Neural Networks (CNNs), in medical image analysis has seen significant advancements in recent years. Several studies have demonstrated the potential of CNN-based systems for automated skin lesion detection and classification.

Esteva et al. (2017) made significant strides in dermatology by employing deep learning techniques to develop a CNN model capable of achieving dermatologist-level accuracy in skin cancer classification. Using a dataset of over 129,000 clinical images, their model demonstrated remarkable effectiveness in distinguishing between malignant and benign lesions. Similarly, Codella et al. (2017) investigated the potential of ensemble deep learning approaches for melanoma detection. By integrating multiple CNN architectures, their system enhanced classification accuracy and achieved state-of-the-art performance on dermoscopic image datasets..

Tschandl et al. (2018) introduced the HAM10000 dataset, a comprehensive collection of 10,000 dermoscopic images representing various skin lesion types. This dataset has become a standard benchmark for evaluating deep learning models in dermatology research, enabling the training of more generalized and robust systems [3]. Studies leveraging this dataset have reported high accuracy in classifying lesion types, underscoring the effectiveness of CNNs in ectracting meaningful heatures from complex medical images.

Advancements in preprocessing techniques have also contributed to improved model performance. Harangi (2018) proposed a hybrid deep learning approach that combined traditional image processing with CNNs to enhance feature extraction, particularly for small datasets. This method addressed issues such as overfitting and dataset imbalance, which are common challenges in medical image analysis [4].

Recently, Li et al. (2021) explored the use of transfer learning to improve skin lesion classification. By utilizing pre-trained CNN architectures such as ResNet and EfficientNet, their approach achieved high classification accuracy with reduced training times, making it a viable solution for real-world clinical applications [5]. Furthermore, augmentation techniques such as rotation, zooming, and color adjustments have been widely adopted to enhance model generalization by artificially expanding the training data [6].

Despite these advancements, challenges remain, including handling imbalance datasets, achieving high specificity for rare lesion types, and ensuring model interpretability for clinical acceptance. Addressing these limitations continues to be a focus of research in the field, with the aim of developing reliable and deployable systems to support dermatologists in clinical settings.

Problem Statement : Skin cancer, particularly melanoma, poses a serious health challenge due to its high mortality rate if not detected early. Traditional diagnosis relies on subjective and time-intensive assessments by dermatologists, often inaccessible in resource-limited areas. Variability in lesion appearance and dataset imbalances further complicate accurate classification. This study proposes a CNN based system for automated skin lesion detection and classification to enhance diagnostic accuracy and accessibility

III. Methodology

A. Data Acquisition and Preparation

In this project the dataset is most important element for developing the algorithms, models, and systems. It contains approximately 25,000 color images from Kaggle dataset, each highlights the specific skin condition. As shown in the figure 1, the dataset includes examples of eight common skin disorder: dermatofibroma, melanocytic nevus, melanoma, squamous cell carcinoma, actinic keratosis, basal cell carcinoma, benign keratosis, and vascular lesion. The images shows the skin lesions from various body parts, making the dataset diverse and comprehensive. After importing the data, the images were divided into subsets for training, validation, and testing: 80% for training, 10% for validation, and 10% for testing.

B. Image Preprocessing

Image pre-processing is an important step to improve the image quality and to prepare the data for analysis. Raw images often include artifacts like hair, air bubbles, or noise that can make analysis less accurate. The pre-processing phase usually involves the several key tasks. One task is resizing the images to 244x244 pixel to match the input requirements of convolutional neural network(CNN), while keeping important feature. Another critical task is pixel normalization, where intensity valuesare scaled from 0 to 255 down to the range of 0 to 1 by diving each value by 255. This scaling helps in making the data consistent and also allows the model to process it more efficiently.

C. Model Development

The model architecture i.e Convolutional neural network has several layers to extract features. The layers are :

First layer is the Convolutional and Pooling layers , which form the backbone of the model. Convolutional layer extract features such as edges, textures and patterns by applying filters to the input image. Rectified linear unit activation function is used to add non-linearity and to learn complex relationships. Then the Pooling layer particularly max-Pooling to reduce the size of the feature maps.

Second layer is flattened layer which takes the output of first layer as input . Input is now flattened into a onedimensional array. The output of the flattened layer is fed to a fully connected layer. The fully connected layers extract features to identify patterns and enhance the prediction accuracy.

The last layer is the output layer. This layer consists of numbers of neurons corresponding to the target class such as seven types of skin lesions. In this stage we use softmax activation function to generally probability distribution and predicts the final classification..

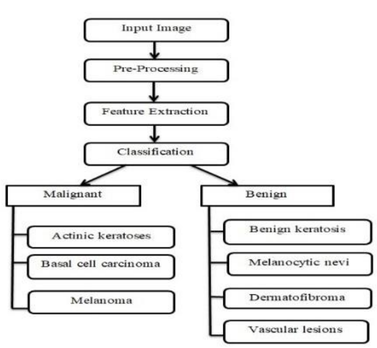
D. Model Training

The data is divided into three categories: training, validation, and testing, with an 80:10:10 distribution. Again, many data enhancement methods are used. These include horizontal and vertical flipping, random rotation, scaling, and image comparison; all designed to diversify the training material and increase the robustness of the model when faced with new patterns

For the loss function, categorical cross entropy is chosen due to its advantages in solving multi-class problems. To improve the training process, Adam optimizer is used, which uses adaptive learning rate. The model started with a learning factor of 0.001 and the learning rate decreased the learning factor by 0.1 times after 5 consecutive epochs without improving the recognition loss. Training was performed more than 50 times with a batch size of 32 and the training process was stopped early to avoid the performance being not good enough for 10 consecutive times.

CNN models are built and trained using TensorFlow and Keras libraries, providing a flexible and flexible model for testing models. Training is done by GPU-enabled systems to speed up the process and support hyperparameter tuning. Throughout the training process, various performance metrics are continuously evaluated to evaluate the model's performance on current data, including precision, accuracy, recall, and F1 score.

The training method of convolutional neural network (CNN) is the main method of skin detection and classification. Initially, a pre-built model using ImageNet dataset formed the basis of transform learning. This enables the network to be fine-tuned to adapt to the specific task of skin lesion classification while taking advantage of the feature extraction capabilities of pre-learning weights.



E. Evaluation and deployment

The trained model was tested on a separate dataset to evaluate its performance and determine its ability to generalize to previously unseen data. For deployment, the model was integrated into a web application using the Flask framework. This setup facilitates real-time predictions by enabling seamless communication between the model and the user interface. Uploaded images are processed through the prediction pipeline, ensuring quick and efficient responses, making the system suitable for use in clinical environments..

Fig 1: Methodology flow diagram

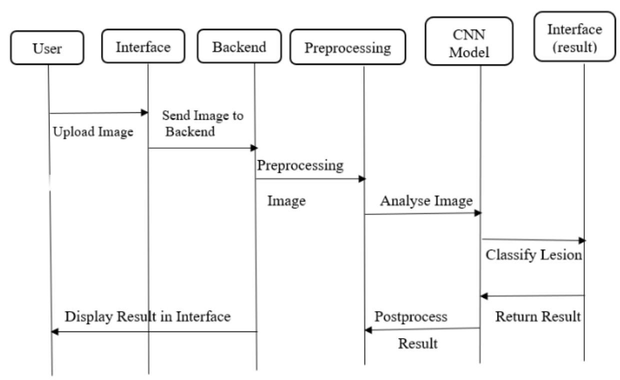


Fig 3: Sequence diagram

IV. Result Analysis

This research applies machine learning and convolutional neural network (CNN) techniques to effectively classify skin lesion images. The experiments were conducted using the HAM10000 dataset, and the performance of both machine learning models and the customized CNN was evaluated based on accuracy, precision, recall, and F1-score. Prior to training and testing, the images underwent pre-processing steps that involved feature and target value extraction, as well as data augmentation to increase the dataset's diversity. The results revealed that the customized CNN achieved a remarkable accuracy of 95.18%, surpassing the performance of the machine learning algorithms used in the study. This highlights the superior classification capability of the proposed CNN model on the HAM10000 dataset. Furthermore, comparisons with recent studies utilizing the same dataset showed that this approach provided improved accuracy while reducing errors and loss. Future research could focus on refining the CNN model by optimizing hyperparameters, including the configuration and types of layers, and exploring different pre-trained models. Additionally, efforts could be directed towards advanced image segmentation techniques and the development of real-time systems for skin lesion classification, aimed at improving accuracy and reducing processing time.

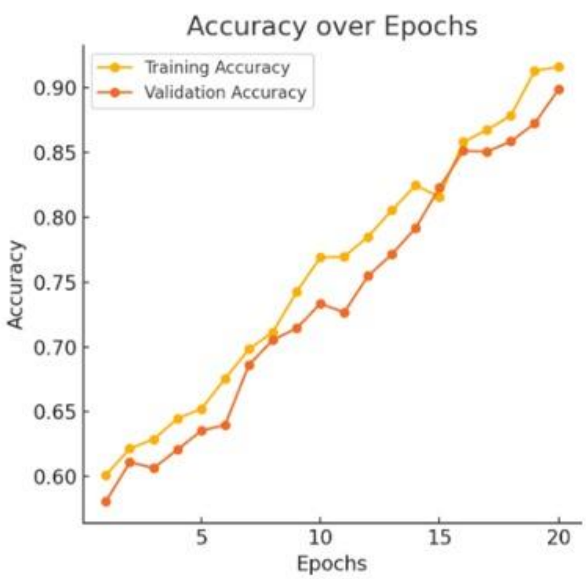


Fig 4: Training and Validation accuracy over the epochs

Training and Validation Accuracy: This exhibit a steady improvement over the epochs, reflecting the model's effective learning from the dataset. The small gap between the two metrics indicates that the model generalizes well to unseen data and does not show substantial evidence of overfitting.

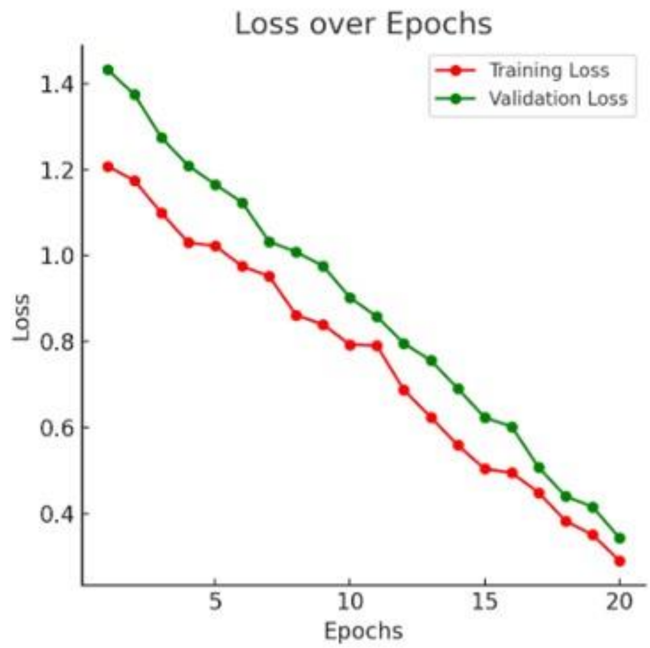


Fig 5 : Training and Validation Loss

Training and validation Loss: This progressively decrease over the epochs, demonstrating the model's ability to learn effectively. The relatively small difference between training and validation losses indicates good generalization to unseen data, with no significant signs of overfitting.

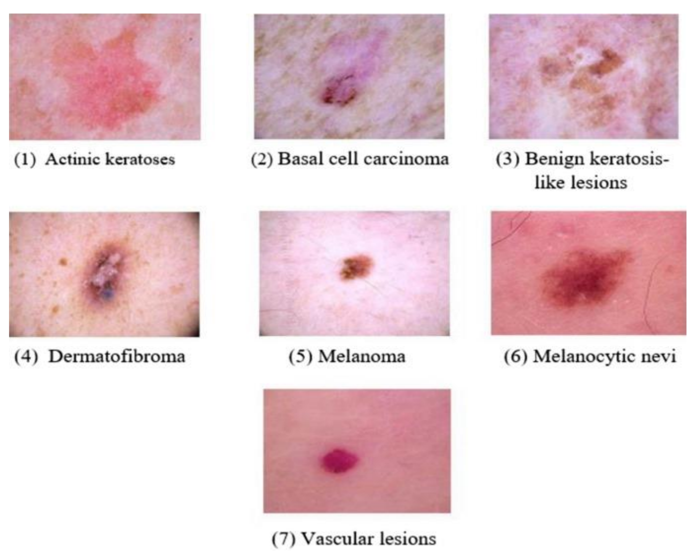


Fig 7 : Figure shows seven types of Skin Lesion

V. Conclusion

This research utilizes machine learning and convolutional neural network (CNN) techniques for the effective classification of skin lesion images. The experiments were carried out using the HAM10000 dataset, and the performance of both machine learning methods and the customized CNN model was assessed based on metrics such as accuracy, precision, recall, and F1-score. Before the training and testing stages, the images underwent pre-processing steps that included feature and target value extraction, along with data augmentation to enhance the dataset's diversity. The findings indicated that the customized CNN achieved a notable accuracy of 95.18%, outperforming the machine learning algorithms employed in the study. This demonstrates the exceptional classification performance of the proposed CNN model on the HAM10000 dataset. Additionally, comparisons with recent studies using the same dataset revealed that the approach delivered improved accuracy while minimizing errors and loss. Future work could focus on further refining the CNN model by optimizing hyperparameters, including the configuration and type of layers, as well as exploring alternative pre-trained models. Moreover, additional efforts could emphasize advanced image segmentation techniques and the development of real-time systems for skin lesion categorization to enhance accuracy and decrease processing time.

VI. References

[1] Taye Girma Debelee- “Skin Lesion Classification and Detection Using Machine Learning Techniques (2023)”.

[2] H M Layes Delower, Tain Mohammad- “Skin Lesion Detection and Classification using Machine Learning: A Comprehensive Approach for Accurate Diagnosis and Treatment (2023)”.

[3] Zahra Mirikharajia, Kumar Abhisheka, Alceu Bissotoc – “A survey on Deep Learning for Skin Lesion Segmentation (2023)”.

[4] Raghav Agarwal and Deepthi Godvarthi – “Skin Disease Classification Using CNN Algorithms (2023)”.

[5] Kmail Dimililer and Boran Sekeroglu – “CNN-based transfer learning model skin lesion diagnosis (2023)”.

[6]”Improving Skin Lesion Diagnosis: Hybrid Blur Detection for Accurate Dermatological Image Analysis” Bhanurangarao mudedla, Mahaveerakannan R (2024).

[7] Jwan Najeeb Saeed, Subhi R. M. Zeebaree-“Skin Lesions Classification Based on Deep Convolutional Neural Networks Architectures (2022)”.

[8] Pranali Jadhav, Pooja Dalvi, Vikas Gaidhani, Navnath Borkar, S. R. Agarwal-“Skin Lesion Classification Using Deep Learning (2023)”.

[9] Ankush Kumar, Dr. Gaurav Gupta-“Skin Lesion Classification” or “Automated Melanoma Detection (2022)”.

[10] R. p. Ram Kumar, Racha Varun, Jageer Sreeshwan, Kondroju Arun Kumar, Upasana Rana, A. Rajyalakshmi-“Feasible Skin Lesion Detection using CNN (2023)”.

[11] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S., Dermatologist-level classification of skin cancer with deep neural networks (2022).

[12] K. M., M.A. Kassem, and M. M Foaud,”Skin Cancer Classification using Deep Learning and Transfer Learning Hosny,” (2023).

[13] A. Romero Lopez, X. Giro-i-Nieto, J.Burdick and O. Marques, “Skin lesion classification from dermoscopic images using deep learning techniques” (2022).

[14] M. A. Kassem, K. M. Hosny, and M. M. Fouad, “Skin Lesions Classification Into Eight Classes for ISIC 2021 Using Deep Convolutional Neural Network and Transfer Learning” (2021).

[15] Bhadoria R.K., Biswas S. (2024) A Model for Classification of Skin Disease Using Pretrained Convolutional Neural Network. In: Mandal j., Mukhopadhyay S. (eds) Proceedings of the Global AI Congress (2024).

[16] Md. Aminur Rab Ratul, Mohammad Hamed Mozaffari, Enea Parimbelli, WonSook Lee, "Atrous Convolution with Transfer Learning for Skin Lesions Classification”.

[17] J. Velasco, C. Pascion, J, Wilmar Alberio, Jonathan Apuang, John Stephen Cruz, Mark Angelo Gomez, Benjamin Jr. Molina, Lyndon Tuala, August Thio-ac, and Romeo Jr. Jorda, “A Smartphone-Based Skin Disease Classification Using MobileNet CNN” (2024)

[18] K.1 Polat, and K. onur Koc, “Detection of the skin Diseases from Dermoscopy Image Using the combination of Convolutional Neural Network and One-versus-All” (2021).

[19] K. Sujay Rao, P.Suresh Yelkar, O. Narayan Pise, and Dr. Swapna Borde, "Skin Disease Detection using

Machine Learning" (2024).

[20] C.Yu Zhu, Yu-Kun Wang, Hai-Peng Chen, Kun-Lun Gao, Chang Shu, Jun-Cheng Wang, Li-Feng Yan, Yi-Guang Yang, Feng-Ying Xie, and Jie Liu, "A Deep Learning Based Framework for Diagnosing Multiple Skin Diseases in a Clinical Environment" (2023).