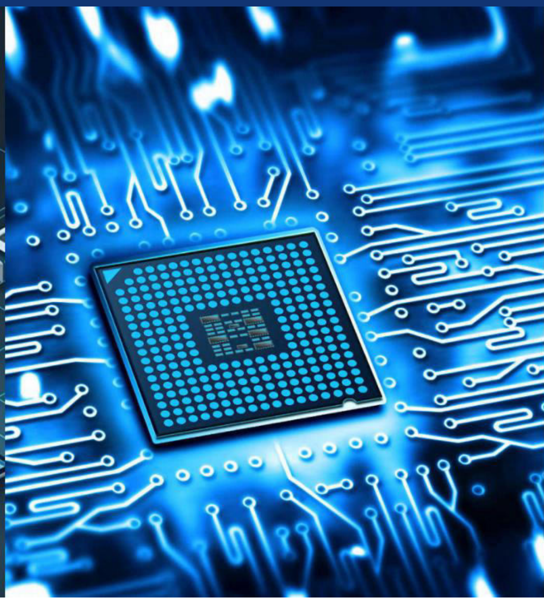
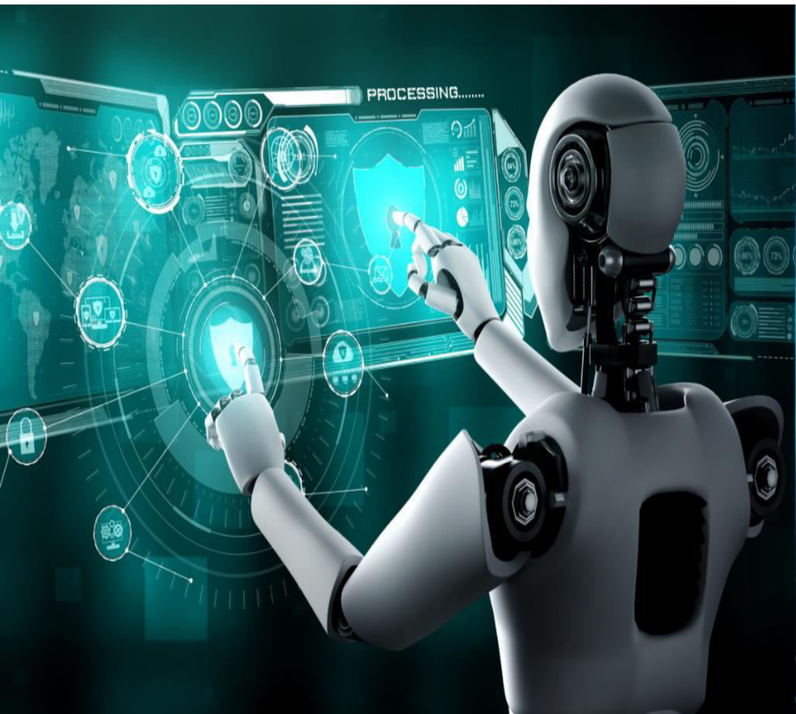
ISSN(O): 2320-9801
ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.771

Volume 13, Issue 5, May 2025

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)
(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Skin Disease Detection System using CNN

Nischal L Sahukar, Rahul Paul, Rakesh M S, Revathi N, Prof. Ranjitha H S

UG Student, Dept. of C.S., Bapuji Institute of Engineering and Technology, Davanagere, Karnataka, India

Assistant Professor, Dept. of C.S., Bapuji Institute of Engineering and Technology, Davanagere, Karnataka, India

ABSTRACT: Skin diseases represent a significant portion of global health issues, ranging from common infections to potentially life-threatening conditions. Timely and accurate diagnosis plays a crucial role in effective treatment and management. However, traditional diagnostic methods are often limited by human expertise and variability in clinical assessments. This project presents a **Skin Disease Detection System using CNN** a novel deep learning architecture that utilizes attention mechanisms to analyze and classify dermatoscopic images with high precision. The system is trained using the **HAM10000 dataset**, a widely recognized collection of annotated skin lesion images encompassing multiple disease classes. Vision Transformers divide images into patches and apply self-attention to capture global contextual information. This approach significantly improves the model's ability to distinguish between subtle variations in skin lesions.

KEYWORDS: Skin Disease Detection, Deep Learning, HAM10000 Dataset, Image Classification, Medical Imaging, Artificial Intelligence, Dermatology.

I. INTRODUCTION

This project presents a skin disease classification system using a custom Convolutional Neural Network (CNN) implemented in PyTorch, trained on the HAM10000 dataset, which consists of dermoscopic images categorized into seven types of skin lesions: melanoma (mel), melanocytic nevi (nv), benign keratosis-like lesions (bkl), basal cell carcinoma (bcc), actinic keratoses (akiec), vascular lesions (vasc), and dermatofibroma (df). The dataset is preprocessed by mapping textual labels to numerical values and loading images from two separate folders. To ensure robust training, data augmentation techniques such as rotation, horizontal flipping, and color jittering are applied to the training set. The data is split into training and validation sets using a stratified approach to maintain class distribution. A simple CNN architecture is designed with three convolutional layers followed by adaptive pooling and a fully connected classifier. To address class imbalance, weighted cross-entropy loss is used, where weights are inversely proportional to class frequencies. The model is trained over multiple epochs with accuracy monitored on both training and validation sets. Post-training, performance is evaluated using a confusion matrix and visualizations of predictions across all classes. This implementation offers a foundational framework for skin lesion classification and can be extended for more complex models such as EfficientNet or Vision Transformers.

II. RELATED WORK

In recent years, deep learning has shown significant promise in the field of medical image analysis, particularly for skin lesion classification. Convolutional Neural Networks (CNNs) have been widely used due to their ability to learn hierarchical features directly from raw images, eliminating the need for manual feature extraction. Several studies have utilized popular architectures such as ResNet, VGGNet, and InceptionNet for classifying dermoscopic images, achieving performance comparable to that of expert dermatologists. The ISIC (International Skin Imaging Collaboration) challenges have further fueled research by providing standardized datasets and evaluation protocols. Many approaches have also addressed the issue of class imbalance in skin cancer datasets through techniques such as oversampling, synthetic data generation (e.g., using GANs), and weighted loss functions. Transfer learning using pretrained models on ImageNet has proven effective in improving accuracy with limited medical data. More recently, hybrid models combining CNNs with attention mechanisms or Vision Transformers (ViTs) have emerged, demonstrating improved performance by focusing on salient regions of skin lesions. This project builds on these developments by implementing a custom CNN model with class weighting to handle imbalance, aiming to provide a baseline for future enhancements using more advanced architectures and techniques.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

1. Data Loading & Preprocessing - Loaded the HAM10000 dataset and mapped skin lesion labels to numerical values.
2. Data Splitting - Used StratifiedShuffleSplit to divide data into training and validation sets while maintaining class distribution.
3. Data Augmentation - Applied transformations such as resizing, random horizontal flipping, rotation, and color jittering to enhance training diversity.
4. Dataset Class Definition - Created a custom SkinCancerDataset class to load and preprocess images dynamically during training and evaluation.
5. CNN Model Architecture - Built a simple CNN with three convolutional layers, ReLU activations, max pooling, adaptive average pooling, and a final fully connected layer.
6. Handling Class Imbalance - Computed inverse class frequencies and applied them as weights in the cross-entropy loss function to reduce bias toward majority classes.
7. Model Training - Trained the model using the **Adam optimizer** with a learning rate of 0.0001 and tracked training/validation accuracy over epochs.
8. Performance Evaluation - Evaluated the model using a confusion matrix, classification accuracy, and prediction visualizations.
9. Visualization of Predictions - Displayed sample predictions for each class to interpret model behavior and validate per-class recognition.

III. SIMULATION RESULTS

The simulation results of the proposed CNN model on the HAM10000 dataset demonstrate promising performance in classifying seven types of skin lesions. Throughout training, the model showed consistent improvements in both training and validation accuracy, indicating effective learning and generalization. The use of data augmentation and stratified splitting contributed to improved robustness and class balance during evaluation. The confusion matrix revealed that the model performed particularly well on majority classes such as nv and mel, while occasional misclassifications were observed in underrepresented classes like vasc and df. To mitigate class imbalance, class weights were applied in the loss function, which enhanced the model's sensitivity to minority classes. Visualizations of predictions across each class further confirmed the model's ability to distinguish between different lesion types. Overall, the lightweight architecture enabled efficient training and testing, making it suitable for rapid experimentation and further development.

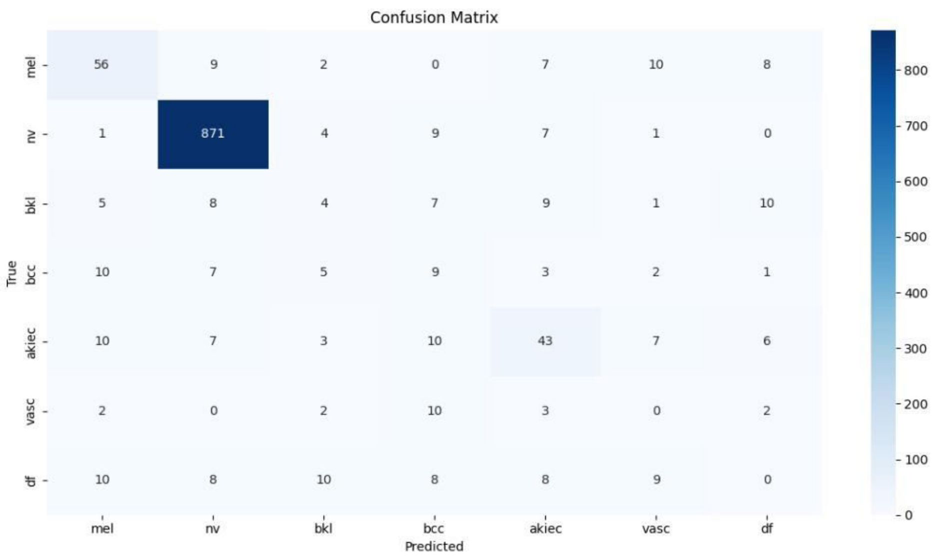


Figure 1 Confusion matrix

Figure 2 Prediction

IV. CONCLUSION

In conclusion, this project demonstrates the effectiveness of a custom Convolutional Neural Network (CNN) in classifying skin lesions using the HAM10000 dataset. By incorporating data augmentation, class balancing through weighted loss, and stratified data splitting, the model was trained to recognize seven types of skin conditions with improved generalization. While the current model provides a solid baseline, there is room for further enhancement through the use of deeper architectures, transfer learning with pretrained models like EfficientNet, or integration of attention mechanisms and Vision Transformers. This work highlights the potential of deep learning in assisting dermatologists with early and accurate diagnosis of skin diseases and lays the groundwork for more advanced research in automated medical image analysis.

REFERENCES

1. **Zhang, X., & Zhao, Y. (2021).** Deep Learning for Skin Disease Detection: A Survey. Journal of Healthcare Engineering, 2021. **DOI: 10.1155/2021/8875461.** This paper surveys various deep learning techniques, including convolutional neural networks (CNNs)
2. **Esteva, A., Kuprel, B., & Novoa, R. A. (2017).** Dermatologist-level classification of skin cancer with deep neural networks. Nature, **542(7639), 115-118. DOI: 10.1038/nature21056**. This landmark study presents the use of deep neural networks for classifying skin Cancer.
3. **Codella, N. C. F., et al. (2018).** Skin cancer metastasis detection using deep learning algorithms. Proceedings of the International Conference on Medical Image Computing (MICCAI),2018.
4. **Jha, D., & Sharma, R. (2020).** Convolutional Neural Networks in Medical Image Analysis: A Review. Health Information Science and Systems
5. **Divya Kodi,** Swathi Chundru, “Unlocking New Possibilities: How Advanced API Integration Enhances Green Innovation and Equity,” in Advancing Social Equity Through Accessible Green Innovation, IGI Global, USA, pp. 437-460, 2025.
6. **Haenssle, H. A., et al. (2018).** Artificial Intelligence for Dermatology: What Can We Expect? The Lancet Digital Health, 1**(2), e57–e63. DOI: 10.1016/S25897500(19)30011-4.**

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING