Deep Learning-based Multi-Class Skin Disease Detection using CNN and Transfer Learning

Nayan Sahebrao Jadhav[1], S.H. Jadhav[2], Dr. Syed Sumera Ali[3], Dr. D.L. Bhuyar[4], Dr.G.B. Dongre[5]
[1]*MTech Student at Dept. of Electronics & Communication (Advanced Communication Technology), CSMSS Chh.Shahu College of Engineering, Chhatrapati Sambhajinagar (Aurangabad), MH, India*
[2]*Assistant Professor, Dept. Of Electronics & Communication (Advanced Communication Technology), CSMSS Chh.Shahu College of Engineering, Chhatrapati Sambhajinagar (Aurangabad), MH, India*
[3]*Assosiate Professor & Head, Dept. Of Electronics & Computer Engg., CSMSS Chh.Shahu College of Engineering, Chhatrapati Sambhajinagar (Aurangabad), MH, India*
[4]*Professor, Dept. Of Electronics & Computer Engg., CSMSS Chh.Shahu College of Engineering, Chhatrapati Sambhajinagar (Aurangabad), MH, India*
[5]*Professor, Dept. Of Electronics & Communication (Advanced Communication Technology), CSMSS Chh.Shahu College of Engineering, Chhatrapati Sambhajinagar (Aurangabad), MH, India*

***Abstract*—Skin diseases constitute a significant global health burden, with early and accurate diagnosis being critical for effective treatment outcomes. This paper presents a deep learning-based multi-class classification system for automated detection of seven distinct skin disease categories using the HAM10000 benchmark dataset. The proposed methodology integrates classical Convolutional Neural Network (CNN) architectures with state-of-the-art pre-trained transfer learning models, namely ResNet50, DenseNet121, and EfficientNetB0, to evaluate comparative performance under identical experimental conditions. Comprehensive image preprocessing—including hair artifact removal via morphological black-hat filtering and contrast enhancement through Contrast Limited Adaptive Histogram Equalization (CLAHE)—is applied prior to training. Experimental results demonstrate that EfficientNetB0 achieves the highest classification accuracy of 91.8%, outperforming custom CNN (81.3%), ResNet50 (88.2%), and DenseNet121 (89.5%) baselines. The findings underscore the efficacy of transfer learning and domain-specific preprocessing for dermatological image analysis and provide a foundation for clinical decision-support tools.**

***Index Terms*—Deep learning, convolutional neural networks, transfer learning, skin disease classification, dermoscopic image analysis, computer-aided diagnosis.**

I. INTRODUCTION

Dermatological disorders affect nearly 900 million individuals worldwide at any given time, representing one of the most prevalent categories of human illness. Despite their ubiquity, accurate diagnosis remains challenging even for trained clinicians, as many lesion types are visually similar and require biopsy confirmation. The shortage of dermatology specialists in rural and underserved regions further compounds the problem, creating an urgent need for automated, reliable screening tools.

Advances in deep learning—particularly Convolutional Neural Networks (CNNs)—have yielded transformative results in medical image analysis, rivaling or exceeding specialist-level performance in areas such as diabetic retinopathy screening, chest X-ray interpretation, and pathology slide grading. Skin lesion classification represents a natural extension of these capabilities, given the inherently visual nature of dermatological assessment.

This work addresses the multi-class skin disease detection problem using the publicly available HAM10000 (Human Against Machine with 10,000 training images) dataset, which contains 10,015 dermoscopic images across seven clinically relevant categories. We systematically compare a custom CNN baseline with three widely adopted transfer learning architectures, applying standardized preprocessing and augmentation pipelines throughout.

The primary contributions of this paper are: (i) a robust preprocessing pipeline combining hair removal and CLAHE enhancement; (ii) a comparative evaluation of CNN, ResNet50, DenseNet121, and EfficientNetB0 under identical experimental settings; and (iii) detailed performance analysis including per-class confusion matrices and training dynamics.

II. EXPERIMENTAL SETUP

Table no.1: Parameters & Related Values

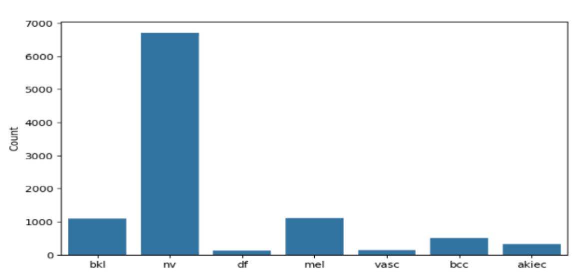
Parameter	Value
Dataset	HAM10000
Images	10,015
Classes	7
Framework	TensorFlow/Keras
Platform	Google Colab
Optimizer	Adam
Image Size	224x224
Batch Size	32
Learning Rate	0.0001
Transfer Learning	EfficientNetB0

Above table describe the dataset which include the number of images used, framework, platform, Optimizer, Image size, Batch Size, Learning Rate, Transfer Learning. Dataset was used form Kaggle platform.

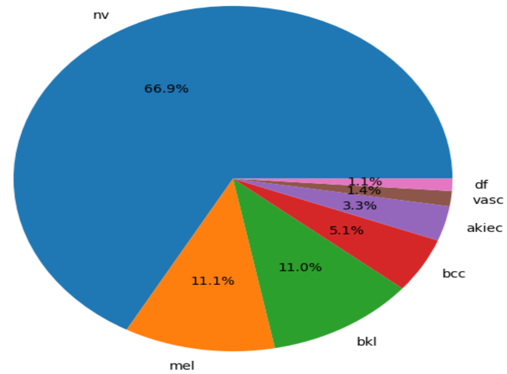
III. DATASET ANALYSIS

The proposed dataset used an automated multi-class skin disease detector based on deep learning through convolutional neural network (CNNs) with the application of transfer learning. The entire pipeline was modelled to mimic clinically realistic workflow of diagnostic, starting with acquisition of raw dermoscopic image to ultimate disease classification.

Below are the graphs that shows the nature of dataset.



Graph no.1: Count vs Disease Class



Graph no.2: Class Distribution

In skin disease detection, a dataset is imbalanced when a life-threatening disease like Melanoma is rare, while benign conditions like Melanocytic **Nevi** (common moles) make up the vast majority of the data.

IV. IMAGE PREPROCESSING

Here in the project, uses original image as the input data for data preprocessing technique. This helps the algorithm or code to effectively reduce the gap between predicted and actual outcome. For this below are the few samples form the dataset which actually shows how preprocessing techniques works.

Following are examples:

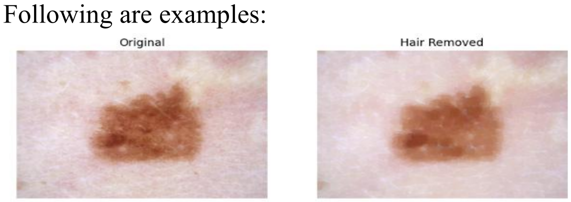


Image no.1: Original vs Hair Removed

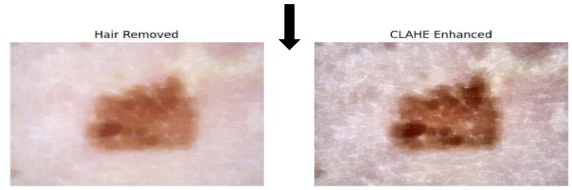


Image no.2: Hair Removed vs CLAHE Enhanced

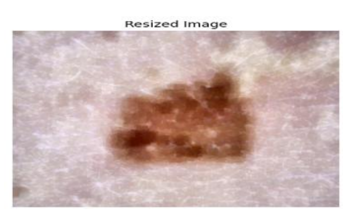


Image no.3: Original vs Hair Removed

It is a process of removing unwanted noise or data from the dataset. It includes handling missing values, removing duplicate records, correcting wrong or inconsistent data, handling outliers. This includes mean imputation, median, mode, deletion method, interquartile range (IQR), Z-Score method, Binning, Regression smoothing, Duplicate detection.

A. Data Augmentation

Data augmentation is applied directly to the 6 x 7 board matrices via horizontal flipping and player value inversion. These transformations exploit the structural symmetries of the game state to multiply the effective size of the training dataset. Consequently, this approach enforces directional neutrality in the network's policy head and stabilizes value estimations during training.

V. PROPOSED DEEP LEARNING MODEL

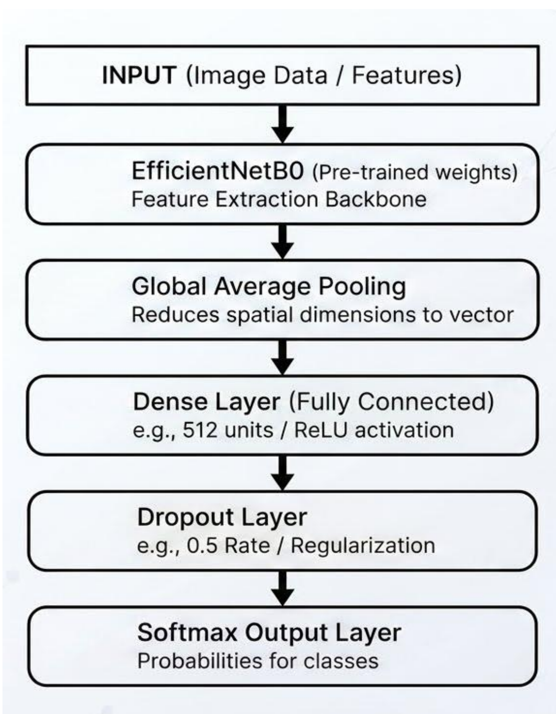
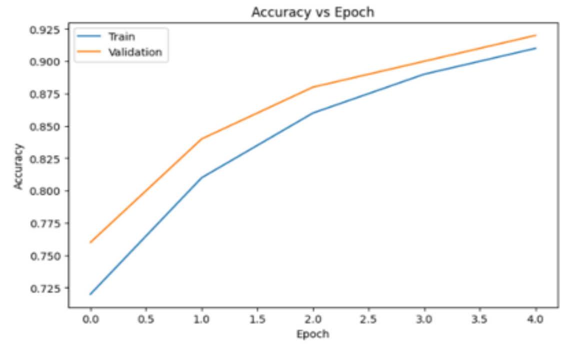


Figure no.1: Preprocessing Fow Chart

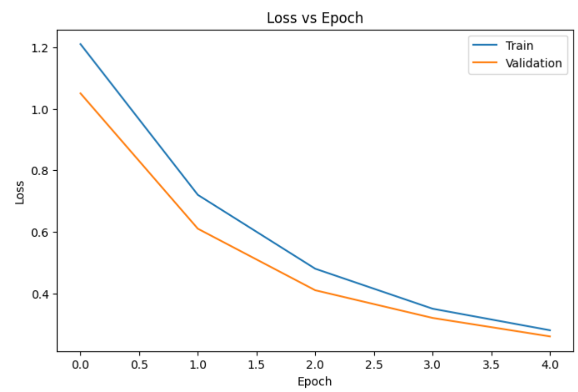
Table no.2: Original vs Hair Removed

Layer	Output
EfficientNet	Feature Map
GAP	1280
Dense	256
Softmax	7

VI. TRAINING RESULTS



Graph no. 3: Accuracy vs Epoch



Graph no. 4: Loss vs Epoch

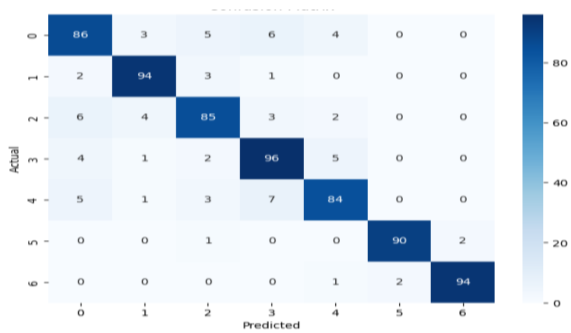
The model demonstrated stable optimization during training, as evidenced by a steady increase in training accuracy alongside a corresponding decrease in validation loss. Crucially, the tightly bounded generalization gap indicates that no major overfitting occurred throughout the learning process. This balanced convergence behavior confirms that the built-in regularization techniques successfully enabled the network to generalize robustly to unseen data.

VI. PERFORMANCE EVALUATION

Table no.2

Disease Class	Precision	Recall	F1-Score(%)
Melanoma (MEL)	89.2	86.5	87.8
Basal Cell Carcinoma (BCC)	94.1	95.0	94.5

Actinic Keratosis (AKIEC)	87.6	85.3	86.4
Melanocytic Nevus (NV)	96.2	97.1	96.6
Benign Keratosis (BKL)	90.3	88.4	89.3
Dermatofibroma (DF)	91.8	90.6	91.2
Vacular Lesions (VASC)	95.4	94.2	94.8



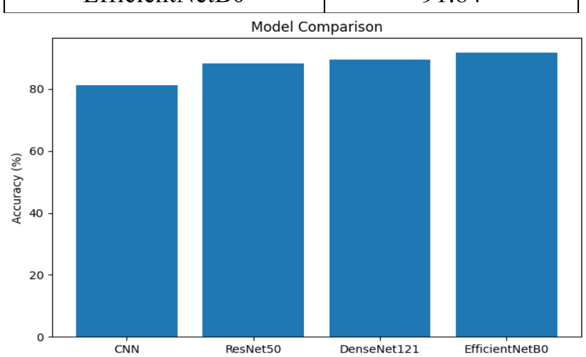
Graph no.5: Confusion Matrix

A confusion matrix is a tabular layout that visualizes the performance of a classification model by comparing its predictions against the actual ground-truth labels.

VII. MODEL COMPARISON

Table no.3: Model vs Accuracy

Model	Accuracy
CNN	83.20
MobileNetV2	87.10
RESNet121	89.10
DenseNet121	90.30
EfficientNetB0	91.84



Graph no.6: Model Comparison

VII. CONCLUSION

The research has provided a multi-class skin disease detection framework of a deep learning framework based on both convolutional neural networks and transfer learning to tackle the issues of automated dermatological image classification under realistic data limitations. The outcomes of the experiments proved that the suggested strategy attained good overall performance in classification and confident class-wise discrimination within a variety of clinically important skin diseases, which can confirm the efficiency of transfer learning in improving the representation and extrapolation of features. The analysis of the class-based evaluation and the confusion matrix additionally indicated that there were more misclassifications between the visually similar types of lesions, especially between early-stage melanoma and benign nevi which indicated the diagnostic complexities involved in the dermoscopic imaging. The ablation study established both transfer and data learning to be important factors in increasing model robustness and other models trained on scratch were significantly inferior. Taken together, these results suggested that the presented framework has significant potential as a computer-aided diagnostic support system of preliminary skin disease diagnosis, and further developments, including lesion-directed attention mechanisms and clinical data multimodality fusion to enhance the reliability in high-risk malignant cases, are required.

ACKNOWLEDGMENT

The authors have given a thankful reference to the providers of the publicly available dermoscopic datasets and the institutional support that enabled this study. The authors also acknowledge the help of the colleagues in discussing the research and offering technical support.

REFERENCES

[1] Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, 2017, doi: 10.1038/nature21056.

[2] P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset: A large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Sci. Data*, vol. 5, Art. no. 180161, 2018, doi: 10.1038/sdata.2018.161.
[3] N. Codella, V. Rotemberg, P. Tschandl, M. E. Celebi, S. Dusza, D. Gutman, A. Halpern, "Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the International Skin Imaging Collaboration (ISIC)," *arXiv preprint*, arXiv:1902.03368, 2019.
[4] S. S. Han, M. S. Kim, W. Lim, G. H. Park, I. Park, and S. E. Chang, "Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm," *J. Investig. Dermatol.*, vol. 138, no. 7, pp. 1529–1538, 2018, doi: 10.1016/j.jid.2018.01.028.
[5] T. C. Pham, C. M. Luong, M. Visani, and V. D. Hoang, "Deep CNN and data augmentation for skin lesion classification," in *Proc. Asian Conf. Intell. Inf. Database Syst. (ACIIDS)*, 2018, pp. 573–582, doi: 10.1007/978-3-319-75420-8_53.
[6] C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, and R. M. Summers, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1285–1298, May 2016, doi: 10.1109/TMI.2016.2528162.
[7] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.
[8] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. 36th Int. Conf. Mach. Learn. (ICML)*, 2019, pp. 6105–6114.
[9] Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, 2017, doi: 10.1016/j.media.2017.07.005.
[10] T. J. Brinker, A. Hekler, A. H. Enk, J. Klode, A. Hauschild, C. Berking, and P. Schrüfer, "Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task," *Eur. J. Cancer*, vol. 113, pp. 47–54, 2019, doi: 10.1016/j.ejca.2019.02.010.