SkinNet-14: a deep learning framework for accurate skin cancer classification using low-resolution dermoscopy images with optimized training time

Abdullah Al Mahmud, Sami Azam, Inam Ullah Khan, Sidratul Montaha, Md.Zahid Hasan

Department of Computer Science and Engineering, Daffodil International University,
Daffodil Smart City, Birulia, Dhaka-1216, Bangladesh

1 Introduction

Skin cancer, including melanoma and non-melanoma types, poses a significant global and health challenge, with 325,000 new melanoma cases and 57,000 deaths reported in 2020. Early detection is critical but often hindered by manual diagnosis challenges and resource constraints. SkinNet-14, a novel deep learning model, leverages a modified Compact Convolutional Transformer (CCT) to classify skin cancer types using low-resolution (32×32 pixel) dermoscopy images. Achieving accuracies of 97.85%, SkinNet-14 reduces training time to 2–8 seconds per epoch, making it ideal for resource-limited clinical settings. This research advances automated diagnostics, enhancing accessibility and efficiency in skin cancer detection.

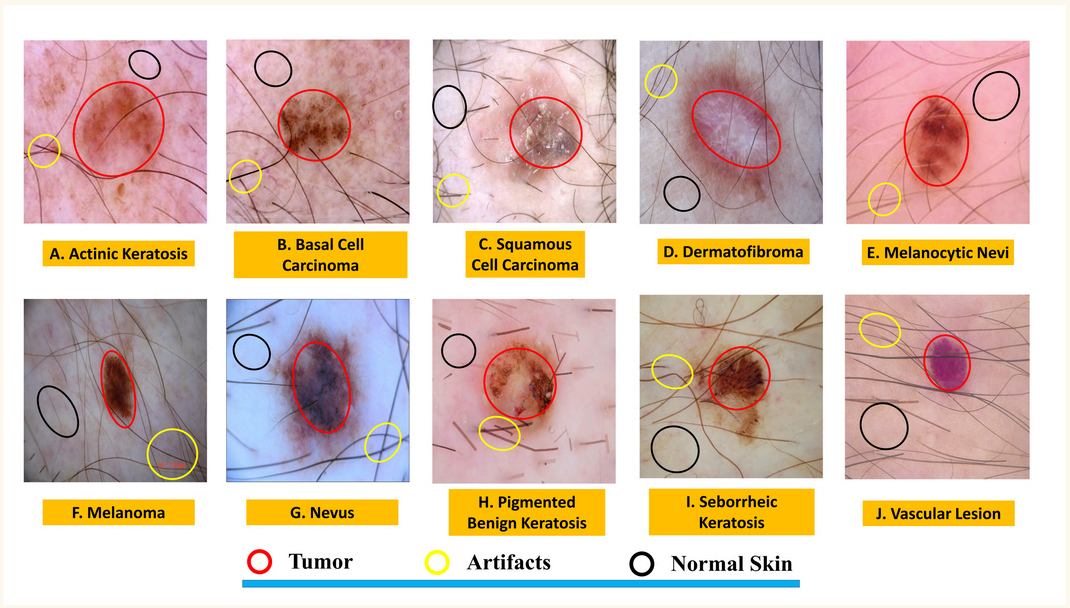

Figure: Cancer lesions of different skin classes

2 Problem Statement

Traditional skin cancer detection methods rely on high-resolution images and computationally intensive models, limiting their use in resource-constrained environments. Class imbalances in medical datasets and lengthy training times further complicate accurate diagnosis.

3 Objectives

1. Develop SkinNet-14 to classify skin cancer using low-resolution dermoscopy images.
2. Minimize computational complexity and training time while maintaining high accuracy.
3. Address dataset imbalances through advanced preprocessing and augmentation techniques.
4. Ensure model stability with reduced training data for real-world applicability.

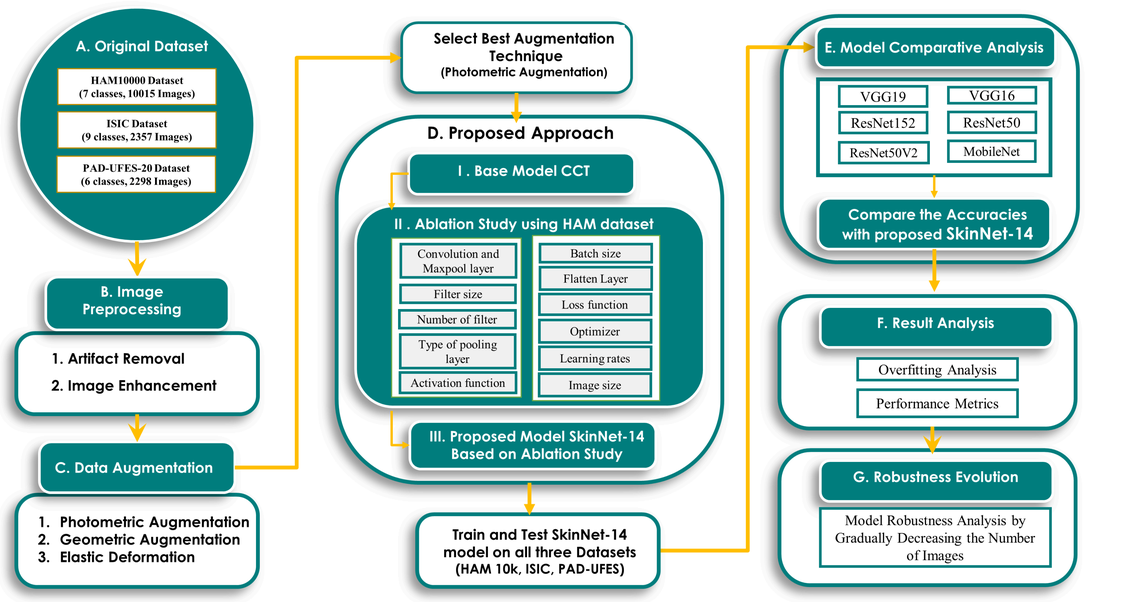

Figure: methodology to classify multiclass skin disease on three datasets

4 Methodology

- Preprocessing: Morphological opening removes artifacts, Non-Local Means Denoising (NLMD) reduces noise, CLAHE enhances contrast, and Gaussian Blur smooths images. Images are resized to 224×224, then downscaled to 32×32.
- Data Augmentation: Photometric augmentation (brightness, contrast, color, sharpness adjustments) balances class distributions, increasing dataset sizes (e.g., HAM10000: 10,015 to 50,075 images).
- Datasets: HAM10000 (10,015 images, 7 classes), ISIC-2019 (2,357 images, 9 classes), PAD-UFES-20 (2,298 images, 6 classes).

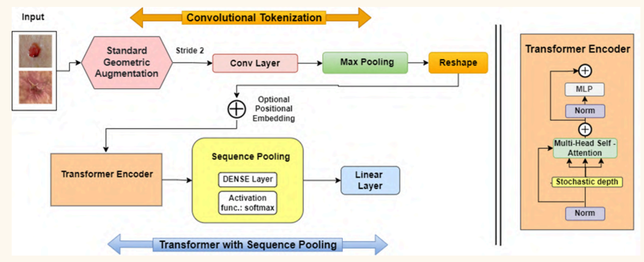

Figure: Structure of Base Model

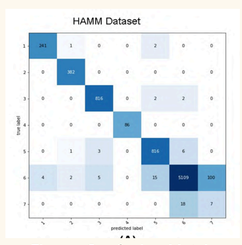

Figure: Confusion Matrix

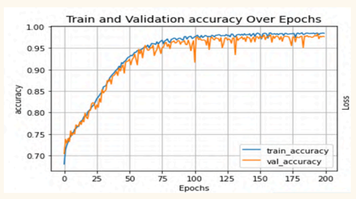

Figure: Accuracy curve of SKINNET-14

5 Results

Model	Parameters	HAM10000DATASET		ISIC DATASET		PAD DATASET	
		Per epoch time	Accuracy	Per epoch time	Accuracy	Per epoch time	Accuracy
VGG19	20026436	65-67s	80.47%	30-34s	70.87%	28-30s	82.97%
VGG16	14716740	65-67s	81.21%	30-34s	71.21%	28-30s	81.38%
ResNet152	58379140	65-67s	65.79%	30-34s	75.79%	28-30s	78.79%
ResNet50	23595908	65-67s	69.27%	30-34s	68.57%	28-30s	72.97%
ResNet50V2	23572996	65-67s	66.25%	30-34s	63.21%	28-30s	77.15%
MobileNet	3232964	65-67s	43.42%	30-34s	49.12%	28-30s	55.48%
SKINNET-14	241861	7-8s	97.85%	2-3s	96.01%	2-3s	98.14%

6 Discussion

- Accuracy: SkinNet-14 achieved 97.85% (HAM10000), 96.01% (ISIC), and 98.14% (PAD-UFES-20), outperforming transfer learning models (e.g., VGG16: 81.21% on HAM10000).
- Training Time: 7–8s/epoch (HAM10000), 1–2s/epoch (ISIC), 2–3s/epoch (PAD), compared to 65–67s/epoch for transfer learning models.
- Stability: Maintains high accuracy with reduced data (e.g., 90.27% with 6,384 HAM10000 images).

Health Informatics Research Lab

7 References

- Hassani, A., et al. (2021). Escaping the Big Data Paradigm with Compact Transformers. CoRR, abs/2104.05704.
- Sabottke, C. F., et al. (2020). The effect of image resolution on deep learning in radiology. Neuroradiology, 62, 1443–1454.

- Made by: Abdullah Al Mahmud; +8801679719544