

SKIN DISEASE DETECTION USING DEEP LEARNING WITH CNN AND ResNet50

T. Kavipriyan

M.Sc Computer Science, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Pallavaram, Tamil Nadu, India

Ms. C.S. Selinchandra, M.C.A., M.Phil., (Ph.D)

Assistant Professor, Department of Computer Science and Information Technology, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Pallavaram, Tamil Nadu, India

Abstract

The exponential rise in skin cancer incidence globally has underscored the urgent need for accurate, automated, and accessible diagnostic tools. Traditional dermatological diagnosis — reliant on visual inspection, dermoscopy, and histopathological biopsy — is inherently subjective, resource-intensive, and inaccessible in underserved settings. This paper presents a deep learning-based Skin Disease Detection system integrating Convolutional Neural Networks (CNN) with the ResNet50 architecture to enable precise, automated classification of skin lesions. The proposed system leverages transfer learning and fine-tuning on large annotated dermoscopic datasets, including the HAM10000 benchmark, to distinguish between benign and malignant lesion types across seven classes: AKIEC, BCC, BKL, DF, MEL, NV, and VASC. The pipeline encompasses image preprocessing, data augmentation, feature extraction via ResNet50, and softmax-based classification through a fully connected layer. Experimental evaluation demonstrates an overall classification accuracy exceeding 90%, with high sensitivity and specificity, and an AUC-ROC score confirming the model's discriminative power. The system is implemented using TensorFlow and MATLAB, with a user-friendly graphical interface enabling healthcare practitioners to upload dermoscopic images and receive real-time diagnostic predictions. Comparative analysis confirms that the proposed CNN-ResNet50 system outperforms traditional image-processing-based classifiers and selected deep learning baselines on identical datasets. This work demonstrates the feasibility of an accurate, scalable, and clinically viable automated skin cancer classification system for deployment in diverse healthcare environments.

Keywords — *Skin Cancer Detection, Convolutional Neural Network, ResNet50, Transfer Learning, Deep Learning, Dermoscopy, Image Classification, HAM10000, Medical Image Analysis, MATLAB.*

1. INTRODUCTION

Skin cancer is one of the most prevalent and potentially fatal forms of cancer worldwide, with incidence rates rising sharply over recent decades. According to the World Health Organization, early-stage diagnosis drastically improves survival rates, yet conventional diagnostic methods remain constrained by limited specialist availability, high inter-observer variability, and the requirement for expensive equipment. Dermatologists currently rely on visual inspection, dermoscopy, and tissue biopsy — methods that are inherently subjective, time-consuming, and inaccessible in rural or resource-limited settings.

The emergence of artificial intelligence (AI) and, specifically, deep learning has introduced transformative possibilities in medical image analysis. Convolutional Neural Networks (CNNs) have demonstrated remarkable performance on image classification tasks, often matching or exceeding specialist-level accuracy by automatically learning hierarchical feature representations from raw pixel data. Architectures such as ResNet50, which address the vanishing gradient problem through residual connections, have proven particularly effective for fine-grained classification tasks involving complex visual patterns such as skin lesions.

This paper presents a Skin Disease Detection system built upon the CNN-ResNet50 architecture. The system is designed to classify dermoscopic skin images into seven clinically relevant categories, providing healthcare practitioners with an automated, reliable second-opinion diagnostic tool. The platform is implemented using TensorFlow and MATLAB, featuring a graphical user interface (GUI) that supports real-time image upload and prediction, making it suitable for deployment in clinical environments.

The following sections provide a comprehensive literature review of related work, a detailed description of the problem statement, the proposed system architecture and methodology, implementation details, experimental results, and a discussion of findings and future directions.

2. PROBLEM STATEMENT

Existing skin cancer classification systems suffer from a number of critical limitations that undermine their clinical utility:

- **High Subjectivity:** Visual inspection by dermatologists is inherently subjective. Diagnostic decisions can vary significantly based on practitioner experience, environmental conditions, and image quality, leading to inconsistent outcomes across clinical settings.
- **Limited Discriminative Power:** Traditional image processing techniques and rule-based classifiers rely on manually engineered features, which often fail to capture the subtle morphological patterns that distinguish malignant from benign lesions, resulting in misclassification.
- **Poor Scalability and Accessibility:** Conventional diagnostic workflows require specialized equipment and trained specialists. In rural and underserved regions, access to dermatological expertise is severely restricted, leaving large populations without timely diagnosis.
- **Lack of Multi-Class Classification:** Many existing automated systems address binary benign-versus-malignant classification, failing to differentiate among the clinically distinct subtypes of skin lesions that require different treatment protocols.
- **Computational Overhead of Prior Models:** Earlier deep learning approaches, such as those based on DenseNet or custom CNNs with dozens of layers, introduce prohibitive computational and training costs that are impractical for standard clinical infrastructure.
- **Interpretability Deficit:** Deep learning models are often perceived as black boxes. Without transparency in predictions, clinicians may be reluctant to integrate automated systems into diagnostic workflows.

3. RELATED WORK

3.1 Deep Learning-Based Skin Cancer Detection

Khamsa et al. (2023) presented a dual-task framework for skin cancer detection and diagnosis using transfer learning with VGG16 for feature extraction and YOLO for object detection, applied to a 3,297-image Kaggle dermoscopic dataset. The model achieved an evaluation accuracy exceeding 83%, demonstrating the viability of multi-task architectures. However, the study's scope was constrained by dataset size and the absence of clinical multi-class benchmarking.

3.2 AI-Assisted Differential Diagnosis Training

Chang et al. (2022) leveraged the InceptionResNetV2 (IRV2) model trained on the HAM10000 database — a widely used benchmark of 10,000 dermoscopic images — to assist dermatology residents in learning differential diagnosis. The system enabled real-time image capture and probability-based prediction, serving as a clinical training aid. While the approach confirms the clinical viability of deep learning in dermatological education, the study did not evaluate deployment performance in live diagnostic settings.

3.3 Deep CNN for Early Skin Cancer Diagnosis

Malaiarasan et al. (2023) proposed a Deep Convolutional Neural Network (DCNN) with 12 nested processing layers specifically designed for skin cancer detection, incorporating Gaussian noise filtering and data augmentation. The study reported superior performance compared to shallow machine learning models and underscored the advantage of architectural depth in feature discrimination. The proposed work builds on this insight by adopting ResNet50 — an architecture that achieves comparable depth while mitigating vanishing gradient issues through residual connections.

3.4 Clinical Information Fusion for Imbalanced Classification

Cao et al. (2022) introduced CICRNet, which combines a Category-Aware (CA) module and a Clinical Information Fusion (CIF) module to address inter-class similarity, intra-class variation, and class imbalance — persistent challenges in skin cancer datasets. Evaluated on the PAD-UFES-20 dataset, the model demonstrated state-of-the-art performance. The insight that clinical metadata can augment purely visual classification directly informs future extensions of the proposed system.

3.5 CNN-Based Multi-Class Classification

Pillai et al. (2023) proposed a 26-layer CNN trained and tested on the HAM10000 dataset, classifying lesions into the same seven categories adopted by this work (AKIEC, BCC, BKL, DF, MEL, NV, VASC). The system achieved a remarkable accuracy of 99.94% through extensive data augmentation. While this result is exceptional, the study used a homogeneous single-dataset evaluation, raising questions about generalisability. The proposed system in this paper adopts a comparable classification taxonomy while emphasizing practical deployability and cross-validation robustness.

4. SYSTEM ARCHITECTURE

The proposed system adopts a three-tier architecture enforcing separation of the user interface, application logic, and deep learning inference layers.

4.1 Presentation Tier

The presentation layer is implemented in MATLAB using the GUIDE (Graphical User Interface Development Environment) toolbox. The GUI provides controls for dermoscopic image upload, preprocessing parameter selection, real-time classification result display, and confidence score visualization. The interface is designed to be intuitive for non-technical clinical users, requiring no knowledge of the underlying machine learning pipeline.

4.2 Application Tier (MATLAB Processing Layer)

The MATLAB application layer orchestrates the end-to-end diagnostic pipeline. Upon image submission via the GUI, the application invokes the preprocessing module (resizing, normalization, augmentation), passes the processed image tensor to the trained ResNet50 model for feature extraction, and forwards extracted features to the fully connected classification head. The system returns the predicted lesion class and associated softmax probability scores to the presentation tier.

4.3 Deep Learning Inference Tier (TensorFlow / ResNet50)

The deep learning inference layer is implemented in TensorFlow (Python 3.x). ResNet50 — pre-trained on ImageNet — is fine-tuned on the HAM10000 dermoscopic dataset. The network architecture consists of: an input layer accepting 224x224x3 tensors; 49 convolutional layers organized into five residual blocks with batch normalisation and ReLU activations; global average pooling; a 1024-unit dense layer with dropout regularization; and a 7-unit softmax output layer corresponding to the seven lesion classes. Trained model weights are exported

and integrated into the MATLAB inference pipeline via TensorFlow Serving or direct Python–MATLAB interoperability.

5. METHODOLOGY

5.1 Dataset and Preprocessing

The primary dataset used for training, validation, and testing is the HAM10000 (Human Against Machine with 10000 training images) benchmark, comprising 10,015 dermoscopic images across seven diagnostic categories: actinic keratoses (AKIEC), basal cell carcinoma (BCC), benign keratosis-like lesions (BKL), dermatofibroma (DF), melanoma (MEL), melanocytic nevi (NV), and vascular lesions (VASC). All images are preprocessed to a uniform resolution of 224x224 pixels. Pixel values are normalized to the range [0, 1]. To address significant class imbalance — particularly the overrepresentation of NV-class images — class-weighted training and online data augmentation are applied, including random horizontal and vertical flips, rotation (up to 30 degrees), brightness and contrast jitter, and cutout regularisation.

5.2 Transfer Learning and Fine-Tuning

ResNet50, pre-trained on ImageNet, is adopted as the feature extraction backbone. The initial convolutional blocks are frozen during a warm-up training phase to preserve low-level feature representations learned on natural images. Subsequently, all layers are unfrozen and fine-tuned end-to-end using a cosine-annealing learning rate schedule with a peak rate of 1e-4 and a minimum of 1e-6. The Adam optimizer is employed with a weight decay of 1e-5. Cross-entropy loss with label smoothing (epsilon = 0.1) is used as the training objective to prevent over-confident predictions.

5.3 Classification Head

Following global average pooling of ResNet50's final convolutional output (2048-dimensional feature vector), a custom classification head is appended: a BatchNormalisation layer; a Dense(1024) layer with ReLU activation; a Dropout(0.5) layer; and a Dense(7) layer with softmax activation. The predicted class corresponds to the argmax of the softmax output, with confidence reported as the maximum softmax probability.

5.4 GUI-Based Inference Workflow

In deployment mode, the clinician uploads a dermoscopic image via the MATLAB GUI. The application preprocesses the image and invokes the TensorFlow inference endpoint. The top predicted class and its softmax confidence score are displayed on the GUI panel, along with a bar chart visualising the probability distribution across all seven lesion classes, enabling clinicians to assess diagnostic confidence at a glance.

5.5 Evaluation Protocol

The dataset is split into training (70%), validation (15%), and test (15%) subsets with stratification to preserve class distributions. Performance is assessed using accuracy, sensitivity (recall per class), specificity, F1-score, and area under the receiver operating characteristic curve (AUC-ROC). The model is trained for a maximum of 100 epochs with early stopping (patience = 10) monitoring validation loss.

6. IMPLEMENTATION

6.1 Technology Stack

Component	Technology
Operating System	Windows 11 (Dev), Ubuntu 20.04 LTS (Deploy)

Programming Language	Python 3.9, MATLAB R2023a
Deep Learning Framework	TensorFlow 2.x / Keras
Model Architecture	ResNet50 (ImageNet pre-trained)
Dataset	HAM10000 (10,015 dermoscopic images)
Image Processing	OpenCV, MATLAB Image Processing Toolbox
GUI Framework	MATLAB GUIDE
Hardware (Training)	NVIDIA GPU, 16GB RAM, 90GB HDD
Evaluation Metrics	Accuracy, Sensitivity, Specificity, AUC-ROC, F1-Score

Table 1: Technology Stack

6.2 Dataset Distribution

Class Code	Lesion Type	Image Count	Split (Train/Val/Test)
MEL	Melanoma	1,113	779/167/167
NV	Melanocytic Nevi	6,705	4,694/1,006/1,005
BCC	Basal Cell Carcinoma	514	360/77/77
AKIEC	Actinic Keratosis	327	229/49/49
BKL	Benign Keratosis	1,099	769/165/165
DF	Dermatofibroma	115	81/17/17
VASC	Vascular Lesions	142	99/22/21

Table 2: HAM10000 Dataset Class Distribution

6.3 Model Training Configuration

Parameter	Value
Base Model	ResNet50 (weights=imagenet)
Input Shape	224 x 224 x 3
Optimizer	Adam (weight_decay=1e-5)
Peak Learning Rate	1e-4 (cosine annealing)
Batch Size	32
Max Epochs	100 (early stopping, patience=10)
Loss Function	Categorical Cross-Entropy (label smoothing=0.1)
Dropout Rate	0.5
Data Augmentation	Flip, Rotation $\pm 30^\circ$, Brightness/Contrast Jitter, Cutout

Table 3: Model Training Configuration

7. RESULTS AND EVALUATION

7.1 Classification Performance

The proposed CNN-ResNet50 system was evaluated on the held-out test set of 1,501 images drawn from all seven lesion classes. The following table summarises the key performance metrics.

Metric	Value
Overall Classification Accuracy	92.3%
Macro-Average Sensitivity (Recall)	89.7%
Macro-Average Specificity	97.1%
Macro-Average F1-Score	90.4%
AUC-ROC (Macro-Average)	0.971
Melanoma (MEL) Sensitivity	88.0%
Melanocytic Nevi (NV) Accuracy	95.1%
Average Inference Time (per image)	0.34 seconds

Table 4: System Performance Results

7.2 Functional Test Results

Test ID	Description	Result
UT-01	Image preprocessing pipeline: resize, normalize, augment	PASS
UT-02	ResNet50 model loading and inference on single image	PASS
UT-03	Softmax output: correct class index and probability distribution	PASS
IT-01	GUI image upload and display confirmed in MATLAB environment	PASS
IT-02	End-to-end pipeline: upload → preprocess → infer → display result	PASS
IT-03	Multi-image batch inference with consistent class assignments	PASS
RT-01	All prior test cases pass after GUI module integration	PASS
WBT-01	Boundary testing: corrupted image, unsupported format, empty input	PASS

Table 5: Functional Test Results

7.3 Comparative Analysis

System	Auto Detect	Multi-Class	Deep Learning	GUI	Acc. >90%
--------	-------------	-------------	---------------	-----	-----------

Manual Dermatologist Review	No	Yes	No	No	Variable
Rule-Based Image Processing	Partial	No	No	No	No
VGG16 + YOLO (Khamsa 2023)	Yes	No	Yes	No	No
DCNN-12 Layer (Malaiarasan 2023)	Yes	Partial	Yes	No	Yes
Proposed CNN-ResNet50 System	Yes	Yes	Yes	Yes	Yes

Table 6: Comparative Analysis

8. DISCUSSION

The proposed CNN-ResNet50 platform demonstrates that automated, deep-learning-based skin lesion classification can achieve clinically meaningful accuracy — 92.3% overall — without the prohibitive computational complexity of larger custom architectures. The adoption of ResNet50 as the backbone strikes a deliberate balance between model capacity and inference efficiency: residual connections mitigate gradient degradation across 50 layers, enabling effective fine-tuning on the relatively small HAM10000 dataset without catastrophic forgetting of ImageNet features.

The macro-average AUC-ROC of 0.971 confirms strong discriminative power across all seven classes, including clinically critical melanoma (MEL) detection, where sensitivity of 88.0% was achieved. This compares favourably to the 83% detection accuracy reported by Khamsa et al. (2023) for a VGG16/YOLO dual-task system, and is competitive with the 99.94% accuracy of Pillai et al.'s (2023) 26-layer custom CNN — which, however, was evaluated only on the HAM10000 training distribution without independent cross-dataset validation.

The average inference time of 0.34 seconds per image confirms the system's suitability for real-time clinical use. The MATLAB GUI enables immediate deployment in environments where Python toolchains may not be maintained, reducing the barrier to clinical adoption.

Limitations of the current implementation include reliance on a single benchmark dataset (HAM10000), which may limit generalisability to out-of-distribution patient populations and imaging devices. The system lacks interpretability beyond class probabilities: clinicians cannot interrogate which lesion regions drove the classification decision. Class imbalance, particularly for the DF and VASC categories, continues to constrain minority-class recall despite augmentation and class weighting.

9. APPLICATIONS

- **Clinical Decision Support:** Dermatology clinics and general practitioners can use the system as an automated second-opinion tool, flagging high-risk lesions for specialist referral and reducing diagnostic delays.
- **Telemedicine and Remote Healthcare:** In rural and underserved areas lacking dermatological specialists, the lightweight MATLAB GUI enables non-specialist healthcare workers to perform preliminary skin cancer screening using dermoscopic or standard camera images.
- **Educational Training Platform:** Medical schools and dermatology training programmes can use the system to expose residents to a diverse range of annotated skin lesion images, accelerating differential diagnosis learning as demonstrated by Chang et al. (2022).

- **Population-Level Screening:** Public health organisations conducting mass skin cancer screening campaigns can deploy the system on mobile devices to triage large populations at low cost, directing limited specialist resources to high-risk individuals.
- **Pharmaceutical and Clinical Research:** Research institutions studying skin lesion prevalence, treatment response, or drug efficacy can use the automated classification system to rapidly annotate large dermoscopic image datasets, reducing manual labelling effort.

10. FUTURE WORK

10.1 Explainability via Gradient-Weighted Class Activation Maps (Grad-CAM)

A primary limitation of the current system is its lack of spatial interpretability. Future work will integrate Grad-CAM visualisation to overlay a heatmap onto input dermoscopic images, highlighting the lesion regions that most strongly influenced the classification decision. This transparency is essential for clinical trust and regulatory compliance of AI diagnostic tools.

10.2 Multi-Dataset Cross-Validation

To improve generalisation, the system will be retrained and evaluated across additional public dermoscopic datasets, including ISIC 2019, PH2, and PAD-UFES-20. Cross-dataset evaluation will expose distributional biases introduced by imaging device variation, patient demographics, and annotation protocols.

10.3 Clinical Metadata Integration

Inspired by the CICRNet architecture of Cao et al. (2022), future versions will incorporate patient metadata — age, anatomical site, gender, and lesion duration — as auxiliary inputs to a clinical information fusion module. This multimodal approach is expected to improve classification accuracy for ambiguous lesion types, particularly AKIEC and BKL, which exhibit significant morphological overlap.

10.4 Mobile and Edge Deployment

To maximise accessibility in resource-constrained environments, the ResNet50 model will be quantised and pruned using TensorFlow Lite to enable deployment on Android and iOS devices. Model compression will be benchmarked against accuracy trade-offs to identify a Pareto-optimal configuration suitable for edge inference without GPU support.

10.5 Tiered Severity Scoring

Future iterations will augment the seven-class softmax output with a derived malignancy risk score on a standardised scale, enabling clinicians to triage patients by urgency and aligning the system output with established clinical decision frameworks such as the ABCDE (Asymmetry, Border, Colour, Diameter, Evolution) melanoma criteria.

11. CONCLUSION

This paper presented a Skin Disease Detection system based on deep learning, integrating Convolutional Neural Networks with the ResNet50 architecture to enable automated, accurate, and accessible classification of dermoscopic skin lesions. The system addresses the key limitations of traditional diagnostic approaches — subjectivity, inconsistency, limited scalability, and restricted accessibility — by embedding state-of-the-art deep learning inference within a clinician-friendly MATLAB graphical interface.

Experimental evaluation on the HAM10000 benchmark dataset demonstrated an overall classification accuracy of 92.3%, a macro-average AUC-ROC of 0.971, and an average inference time of 0.34 seconds per image —

confirming the system's clinical viability across all seven lesion categories. Comparative analysis establishes the proposed system as the only reviewed solution combining automated deep learning-based multi-class skin lesion classification, ResNet50 transfer learning, a deployable MATLAB GUI, and comprehensive multi-metric evaluation in a single coherent platform.

Future work will extend the system with Grad-CAM explainability, multi-dataset cross-validation, clinical metadata fusion, mobile edge deployment, and tiered malignancy risk scoring — advancing toward a fully comprehensive, multimodal, and interpretable clinical decision support system for dermatological healthcare.

REFERENCES

- [1] D. Khamsa, L. Pascal, B. Zakaria, M. Lokman and M. Y. Zakaria, "Skin Cancer Diagnosis and Detection Using Deep Learning," 2023 International Conference on Electrical Engineering and Advanced Technology (ICEEAT), Batna, Algeria, 2023, pp. 1–6, doi: 10.1109/ICEEAT60471.2023.10426611.
- [2] C. H. Chang, W. En Wang, F. Y. Hsu, R. Jhen Chen and H. C. Chang, "AI HAM 10000 Database to Assist Residents in Learning Differential Diagnosis of Skin Cancer," 2022 IEEE 5th Eurasian Conference on Educational Innovation (ECEI), Taipei, Taiwan, 2022, pp. 1–3, doi: 10.1109/ECEI53102.2022.9829465.
- [3] S. Malaivasan, R. Ravi, D. R. Maheswari, C. Y. Rubavathi, M. Ramnath and V. Hemamalini, "Towards Enhanced Deep CNN For Early And Precise Skin Cancer Diagnosis," 2023 International Conference on Networking and Communications (ICNWC), Chennai, India, 2023, pp. 1–7, doi: 10.1109/ICNWC57852.2023.10127521.
- [4] Y. Cao, H. Wang, D. Tian, X. Wang, S. Wu and M. Tian, "CICRNet: Clinical Information and Category Relation Improve Imbalanced Skin Cancer Diagnosis," 2022 7th International Conference on Signal and Image Processing (ICSIP), Suzhou, China, 2022, pp. 596–600, doi: 10.1109/ICSIP55141.2022.9886238.
- [5] R. Pillai, N. Sharma and R. Gupta, "Proposed Convolution Neural Network for Skin Cancer Diagnosis and Classification," 2023 International Conference on Recent Advances in Electrical, Electronics & Digital Healthcare Technologies (REEDCON), New Delhi, India, 2023, pp. 308–312, doi: 10.1109/REEDCON57544.2023.10151029.
- [6] Y. Chen, J. Deng and T. Wang, "Skin Cancer Diagnosis and Medical Service System Based on Deep Learning Models," 2022 3rd International Conference on Electronic Communication and Artificial Intelligence (IWECAL), Zhuhai, China, 2022, pp. 367–371, doi: 10.1109/IWECAL55315.2022.00078.
- [7] N. Darapaneni et al., "Early Detection of Skin Cancer — Solution for Identifying and Defining Skin Cancers using AI," 2022 Interdisciplinary Research in Technology and Management (IRTM), Kolkata, India, 2022, pp. 1–6, doi: 10.1109/IRTM54583.2022.9791676.
- [8] J. Vineeth et al., "Skin Cancer detection using deep learning," 2022 Fourth International Conference on Cognitive Computing and Information Processing (CCIP), Bengaluru, India, 2022, pp. 1–6, doi: 10.1109/CCIP57447.2022.10058685.
- [9] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.
- [10] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778.
- [11] Tschandl, P., Rosendahl, C., & Kittler, H. (2018). The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific Data*, 5, 180161.
- [12] Esteva, A. et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115–118.
- [13] The MathWorks, Inc. (2023). MATLAB R2023a Documentation. Retrieved from <https://www.mathworks.com/help/matlab/>
- [14] TensorFlow Developers. (2024). TensorFlow Documentation. Retrieved from <https://www.tensorflow.org/>
- [15] International Skin Imaging Collaboration. (2023). ISIC Archive. Retrieved from <https://www.isic-archive.com/>