An Optimized Classification Framework for Skin Lesion Detection Using Machine Learning

Shaik Razia
Singapore Institute of Technology, Singapore
razia28sk@gmail.com (corresponding author)

Sivaneasan Balakrishnan
Singapore Institute of Technology, Singapore
sivaneasan@singaporetech.edu.sg

Mohammed Ali Hussain
Sreenidhi Institute of Science and Technology, Hyderabad, India
alihussain.phd@gmail.com

Prasun Chakrabarti
Department of Computer Science and Engineering, Sir Padampat Singhania University, Udaipur Rajasthan, India
drprasun.cse@gmail.com

Received: 28 June 2025 | Revised: 9 July 2025, 13 July 2025, 17 July 2025, 20 July 2025, 28 July 2025, and 4 August 2025 | Accepted: 14 August 2025

 DOI: https://doi.org/10.48084/etasr.13009

ABSTRACT

This study presented a Machine Learning (ML) evaluation for automated skin lesion classification, utilizing three available datasets: DermNet, PH2, and ISIC. The methodology involved preprocessing steps, including image normalization, Gaussian noise filtering, data augmentation through Random Oversampling and SMOTE, and dimensionality reduction using Principal Component Analysis (PCA). Four classical ML classifiers, including Support Vector Machine (SVM), Random Forest (RF), k-Nearest Neighbors (k-NN), and Logistic Regression, were trained and examined. Metrics, such as accuracy, precision, recall, and F1-score, highlighted the variations in classifier effectiveness across datasets. The results demonstrated that the RF model achieved the highest accuracy of 99.3% on the ISIC dataset, while SVM yielded great performance on the DermNet and PH2 datasets, with accuracies of 93.1% and 94.2%, respectively. Future work should focus on incorporating Convolutional Neural Networks (CNNs) and non-visual data.

Keywords-skin lesion detection; machine learning; principal component analysis; class imbalance; classifier performance

I. INTRODUCTION

The early detection and classification of skin lesions is essential in dermatology, as it can significantly improve patient outcomes [1]. Skin lesions can be found in different conditions, ranging from benign to malignant melanoma, each requiring a unique diagnostic and therapeutic approach [2]. The complexity of visual similarity among lesion types and the inter-class variability present substantial challenges for traditional diagnostic methods [3].

The utilization of ML and artificial intelligence has gained a lot of attention in order to mitigate these challenges [4]. ML algorithms have the ability to recognize subtle patterns and tones within skin lesion images more efficiently than human eyes in some cases [5, 6]. These algorithms, particularly those employing deep learning architectures, can autonomously learn hierarchical feature representations, offering a data-driven alternative to feature engineering [7]. Consequently, ML-based approaches are increasingly being adopted to detect skin lesions with accuracy and efficiency in dermatology [8]. This study presents an ML framework for automated skin lesion classification, utilizing three available datasets.

II. METHODOLOGY

This research evaluated the performance of ML classifiers for skin lesion detection [9]. Specifically, several stages were integrated, including data collection, preprocessing, training, evaluation, and visualization, to avoid possible errors [10].

Challenges, such as class imbalance, high-dimensional data, and variability in imaging conditions, were tailored to further ensure reliability [11].

A. Data Collection and Organization

This study utilized three datasets: Dermnet, PH2, and ISIC, each containing various skin lesion images [12-15]. DermNet is a dermatology image database with various clinical skin images for educational and research use. PH2 is a dermoscopic image dataset containing 200 images, specialized in melanoma, atypical, and benign lesions. Additionally, ISIC provides the largest available dataset for skin lesion analysis with thousands of dermoscopic images.

These datasets provided different image quality, lesion types, and class distributions, enhancing the evaluation [16]. Images were pre-organized into class-specific directories, streamlining the labeling and loading processes [17]. Each one was resized to a resolution of 64×64 pixels, for uniformity and compatibility across models [18]. Additionally, metadata, such as lesion type and imaging modality, were incorporated for potential stratified analysis. Figure 1 depicts the flow chart of the proposed model.

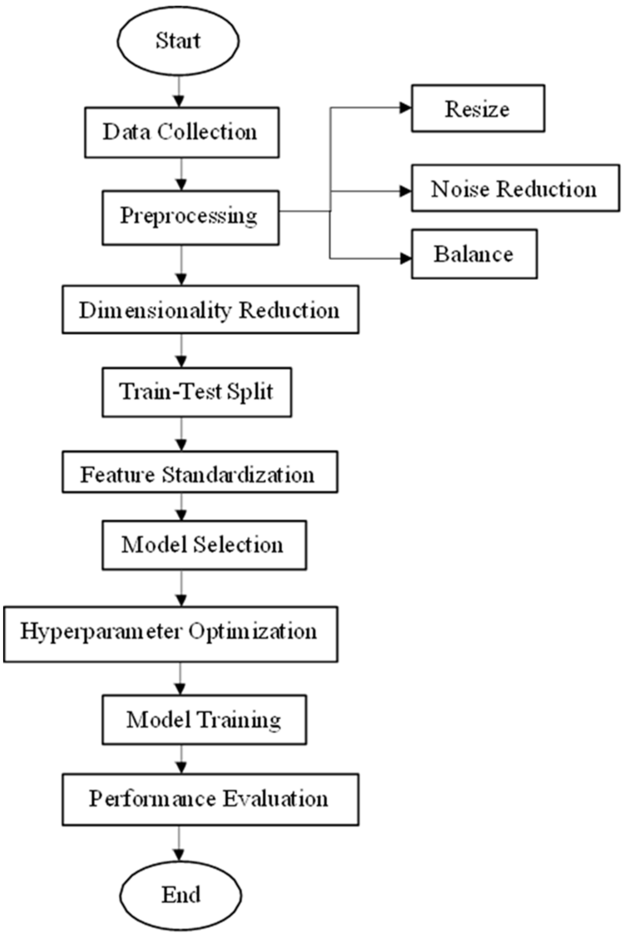

Fig. 1. Flow chart of the proposed model.

B. Preprocessing Pipeline

Preprocessing plays a fundamental role in data quality improvement and in overcoming inherent problems. The whole process comprised the following stages:

- Image Resizing: All images were sized to 64 x 64 pixels for uniformity in input dimensions.
- Noise Reduction: Gaussian filtering was applied to remove image noise while maintaining essential features.
- Class Imbalance Handling: Class imbalance involved random oversampling, which synthetically duplicated underrepresented samples to ensure balanced data distributions across classes. The Synthetic Minority Oversampling Technique (SMOTE) was also considered to generate artificial samples by interpolating existing minority class data points [19].
- Dimensionality Reduction: PCA was deployed to simplify a large dataset into a smaller set while still maintaining significant patterns and trends. In this study, 100 principal components were selected.

C. Classifier Selection and Training

Four ML classifiers were employed in this study:

- SVM: Utilized a Radial Basis Function (RBF) kernel to achieve non-linear separability in high-dimensional feature spaces.
- Random Forest (RF): An ensemble learning technique that aggregates predictions from multiple decision trees, enhancing robustness and reducing overfitting [20].
- k-NN: A distance-based algorithm that classifies instances based on their proximity to labeled data points in the feature space [12].
- Logistic Regression (LR): A baseline linear classifier used for its simplicity and interpretability in understanding the feature contributions [21].

Each classifier underwent hyperparameter optimization through grid search and cross-validation to find the best configuration for each dataset [22]. This approach confirmed that all classifiers would be evaluated under comparable conditions [23]. For k-NN, a range of k from 3 to 15 was used, and it was determined that $k = 5$ yielded the best performance across all datasets. For the SVM, the RBF kernel with a penalty parameter $C = 1.0$ and gamma set to 'scale' were utilized. The RF classifier was configured with 100 estimators, a maximum depth of 10, and bootstrap enabled. LR was trained using the liblinear solver with L2 regularization.

III. RESULTS AND ANALYSIS

Table I presents the performance of SVM, RF, k-NN, and LR across the DermNet, PH2, and ISIC datasets. It was observed that RF exhibited the optimum performance with an accuracy of 99.3% on the ISIC dataset. Figure 2 shows the confusion matrices for the Dermnet dataset.

TABLE I. PERFORMANCE METRICS ACROSS DATASETS AND CLASSIFIERS

Dataset	Classifier	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
DermNet	SVM	93.1	93.0	93.2	93.1
DermNet	RF	95.2	95.0	95.4	95.2
DermNet	k-NN	91.9	91.8	92.0	91.9
DermNet	LR	89.3	89.2	89.4	89.3
PH2	SVM	94.2	94.4	94.7	94.5
PH2	RF	96.1	96.1	96.1	96.1
PH2	k-NN	91.7	92.1	92.3	92.2
PH2	LR	89.7	90.2	90.5	90.4
ISIC	SVM	92.1	91.6	91.8	91.7
ISIC	RF	99.3	99.2	99.3	99.3
ISIC	k-NN	90.9	90.8	91.0	90.9
ISIC	LR	88.5	88.4	88.6	88.5

Figure 3 displays the confusion matrices for the classification performance of SVM, RF, k-NN, and LR models on the PH2 dataset. SVM, RF, and k-NN showed relatively balanced performance with minor misclassifications, while LR exhibited moderate classification accuracy. Similarly, Figure 4 presents the confusion matrices for the classification performance of SVM, RF, k-NN, and LR on the ISIC dataset. SVM demonstrated consistent performance with some misclassifications across classes, while RF showed a balanced classification with slightly fewer errors. k-NN performed well for specific classes but struggled with misclassifications in others, and LR provided moderate but uniform classification accuracy.

These findings highlighted the variability in model performance when applied to different datasets, focusing on the challenges of accurate skin lesion classification.

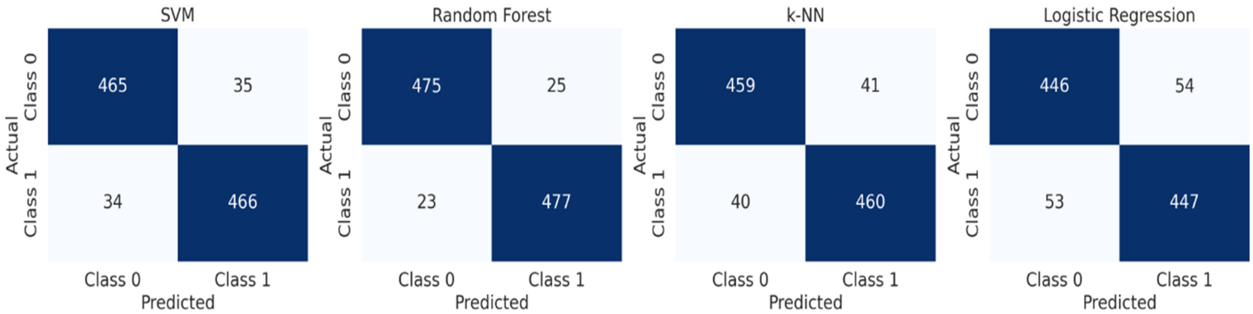

Fig. 2. Confusion matrices on the Dermnet dataset.

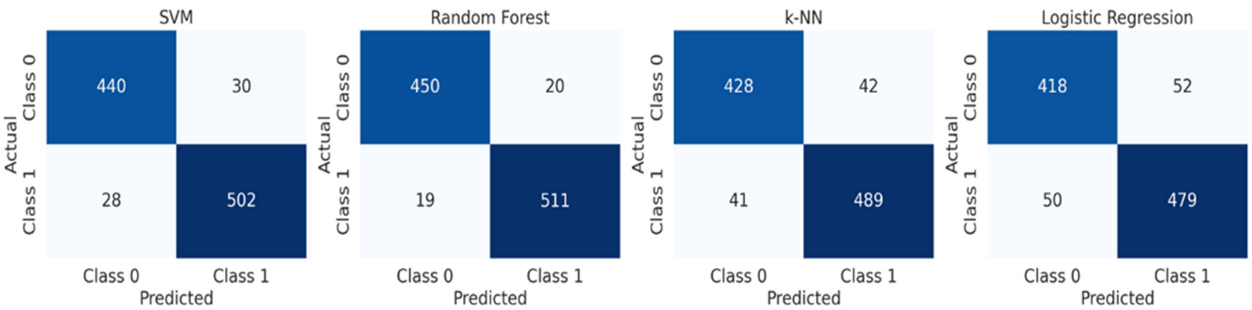

Fig. 3. Confusion Matrices on the PH2 dataset.

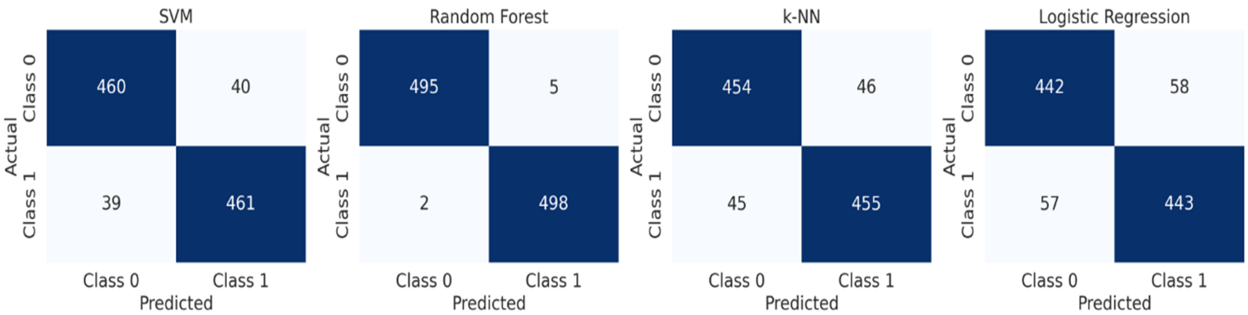

Fig. 4. Confusion matrices on the ISIC dataset.

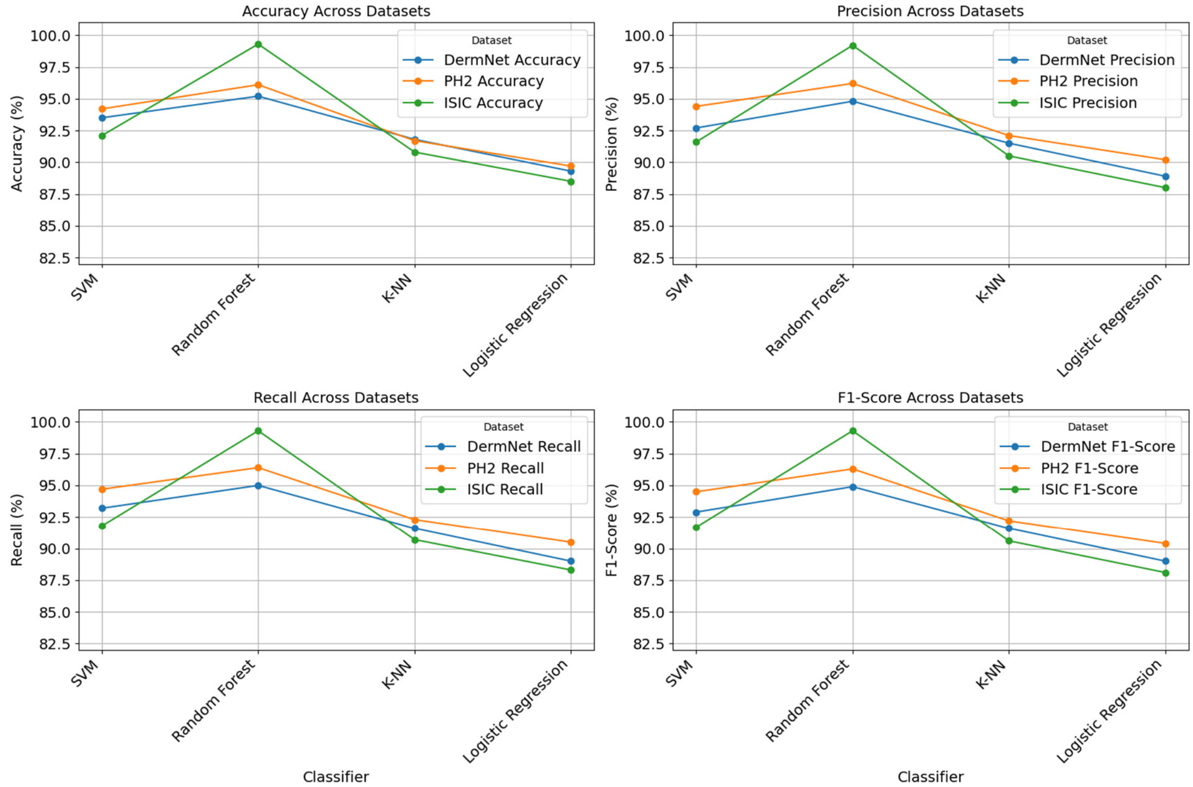

Fig. 5. Output metrics of ML claasifierss across all datasets.

Figure 5 illustrates the output metrics of the ML classifiers across the three ISIC datasets. For DermNet and PH2 datasets, SVM and RF revealed competitive results, while LR and k-NN trailed slightly in overall performance. This visualization provided valuable insights into the classifier behaviors across datasets, helping to identify the optimal model for specific use cases or datasets.

IV. CONCLUSION AND FUTURE SCOPE

This research focused on the examination of different Machine Learning (ML) classifiers on skin lesion detection by using DermNet, PH2, and ISIC datasets. Random Forest classifier was found to have performed the best among all, with an accuracy of 99.3% on the ISIC dataset. Other classifiers like Support Vector Machine (SVM) also worked great, especially on DermNet and ISIC datasets, demonstrating its efficacy in addressing non-linear decision boundaries. On the other hand, Logistic Regression and k-Nearest Neighbors (k-NN) showed a moderate performance, which means that they are not efficient for high-variance datasets.

Future work could rely on the latest Convolutional Neural Networks (CNNs), such as EfficientNet or Vision Transformers (ViT), as well as to examine the addition of non-visual data like patient demographics, medical history, and genetic information to improve diagnostic accuracy.

REFERENCES

[1] G. G. De Angelo, A. G. C. Pacheco, and R. A. Krohling, "Skin lesion segmentation using deep learning for images acquired from smartphones," in *2019 International Joint Conference on Neural Networks (IJCNN)*, Budapest, Hungary, July 2019, https://doi.org/10.1109/IJCNN.2019.8851803.

[2] P. Kaler, S. Kodli, and S. Anakal, "Diagnosis of Skin Cancer Using Machine Learning and Image Processing Techniques," *International Journal of Education and Management Engineering*, vol. 12, no. 5, pp. 38–45, 2022, https://doi.org/10.5815/ijeme.2022.05.05.

[3] N. Mittal, S. Tanwar, and S. K. Khatri, "Identification & enhancement of different skin lesion images by segmentation techniques," in *6th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, Noida, India, Sept. 2017, https://doi.org/10.1109/ICRITO.2017.8342500.

[4] S. N. Hasan, "Accurate Deep Learning Algorithms for Skin Lesion Classification," *International Information and Engineering Technology Association*, vol. 29, no. 4, pp. 1529–1539, Aug. 2024, https://doi.org/10.18280/isi.290426.

[5] S. A. Hanum, A. Dey, and M. A. Kabir, "An Attention-Guided Deep Learning Approach for Classifying 39 Skin Lesion Types." arXiv, Jan. 10, 2025, https://doi.org/10.48550/arXiv.2501.05991.

[6] V. A. Rajendran and S. Shanmugam, "Automated Skin Cancer Detection and Classification using Cat Swarm Optimization with a Deep Learning Model," *Engineering, Technology & Applied Science Research*, vol. 14, no. 1, pp. 12734–12739, Feb. 2024, https://doi.org/10.48084/etasr.6681.

[7] I. Rahman, M. K. Islam, A. N. Chy, and M. A. Azim, "Fusion of Shallow and Deep Features for Classifying Skin Lesions," in *25th International Conference on Computer and Information Technology (ICCIT)*, Cox's Bazar, Bangladesh, Dec. 2022, https://doi.org/10.1109/ICCIT57492.2022.10055219.

[8] S. Gomathi and N. Arunachalam, "Skin Lesion Prediction and Classification Using Innovative Modified Long Short-Term Memory-Based Hybrid Optimization Algorithm," *International Journal of Computational Intelligence Systems*, vol. 17, no. 1, July 2024, Art. no. 186, https://doi.org/10.1007/s44196-024-00599-1.

[9] D. Kourav and A. Kathal, "Skin Lesion Image Segmentation Based on C-Means Clustering Algorithm," *International Journal of Innovative*

Technology and Exploring Engineering (IJITEE), vol. 8, no. 12, pp. 987–990, Oct. 2019, https://doi.org/10.35940/ijitee.K1306.1081219.

[10] A. Upadhyay, A. Chauhan, and D. Kudtarkar, "Skin Lesion Melanoma Detection Using Digital Image Processing," *International Journal of Research and Analytical Reviews (IJRAR)*, vol. 6, no. 1, pp. 44–49, Mar. 2019.

[11] V. Chakkarapani, S. Poornapushpakala, and S. Suresh, "Enhancing Skin Cancer Detection with Multimodal Data Integration: A Combined Approach Using Images and Clinical Notes," *SN Computer Science*, vol. 6, no. 1, Jan. 2025, Art. no. 72, https://doi.org/10.1007/s42979-024-03601-x.

[12] K. Behara, E. Bhero, and J. T. Agee, "An Improved Skin Lesion Classification Using a Hybrid Approach with Active Contour Snake Model and Lightweight Attention-Guided Capsule Networks," *Diagnostics*, vol. 14, no. 6, Mar. 2024, Art. no. 636, https://doi.org/10.3390/diagnostics14060636.

[13] *Dermnet*, Kaggle, https://www.kaggle.com/datasets/shubhamgoel27/dermnet.

[14] *PH2 Dataset*, Dataset Ninja, https://datasetninja.com/ph2.

[15] *ISIC Challenge Datasets*, ISIC Challenge, https://challenge.isic-archive.com/data/?utm_source.

[16] S. Sreena and A. Lijiya, "Skin Lesion Analysis Towards Melanoma Detection," in *2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*, Kannur, India, July 2019, https://doi.org/10.1109/ICICICT46008.2019.8993219.

[17] P. E. Freer, "Skin Lesions," in *Breast Imaging*, Onford, UK: Oxford University Press, 2018.

[18] W. A. Mahdi, S. S. Imam, A. Alotaibi, S. Alhallaf, R. F. Alzhrani, and S. Alshehri, "Formulation and Evaluation of a Silymarin Inclusion Complex-Based Gel for Skin Cancer," *ACS Omega*, vol. 10, no. 3, pp. 3006–3017, Jan. 2025, https://doi.org/10.1021/acsomega.4c09614.

[19] M. Mansilla-Polo, "Should patients undergoing hematopoietic stem cell transplantation undergo screening and monitoring for skin cancer?," *Anais Brasileiros de Dermatologia*, vol. 100, no. 2, pp. 372–373, 2025, https://doi.org/10.1016/j.abd.2024.06.007.

[20] R. Zou, J. Zhang, and Y. Wu, "Skin Lesion Segmentation through Generative Adversarial Networks with Global and Local Semantic Feature Awareness.," *Electronics*, vol. 13, no. 19, Oct. 2024, Art. no. 3853, https://doi.org/10.3390/electronics13193853.

[21] P. Entezari, A. Alaini, H. Mirfazaelian, and Y. Daneshbod, "A cutaneous lesion," *Internal and Emergency Medicine*, vol. 10, no. 7, pp. 879–880, Oct. 2015, https://doi.org/10.1007/s11739-015-1234-4.

[22] D. Zakria *et al.*, "The Role of Image-Guided Superficial Radiation Therapy in the Treatment of Nonmelanoma Skin Cancer," *SKIN The Journal of Cutaneous Medicine*, vol. 9, no. 1, pp. 2042–2054, Jan. 2025, https://doi.org/10.25251/skin.9.1.1.

[23] P. Sahu, S. K. Mohapatra, P. K. Sarangi, J. Mohanty, and P. K. Sarangi, "Detection of non-melanoma skin cancer by deep convolutional neural network and stochastic gradient descent optimization algorithm," *Journal of Mechanics of Continua amd Mathematical Sciences*, vol. 20, no. 1, pp. 59–72, Jan. 2025.