

An Improved Multiclass Skin Cancer Identification Approach with Random Forest Algorithm over HAM10000 Data set using Color Analysis

Sumit Kumar Banerjee¹, Bidisha Bhabani², Mahamuda Sultana³, Suman Bhattacharya⁴

¹Computer Science and Engineering, JIS University, Guru Nanak Institute of Technology, 81, Nilgunj Rd, Jagarata Pally, Deshpriya Nagar, Agarpara; 157/F, Nilgunj Road, Panihati, Sodepur, Kolkata, 700109.

²Computer Science and Engineering, JIS University, 81, Nilgunj Rd, Jagarata Pally, Deshpriya Nagar, Agarpara. 3157/F, Nilgunj Road, Panihati, Sodepur, Kolkata, 700109. 4157/F, Nilgunj Road, Panihati, Sodepur, Kolkata, 700109.

ABSTRACT

Initial diagnosis of melanoma is disapprovingly important, as it significantly improves the patient's chances of survival. Identification of melanoma from dermoscopic images is an inspiring task due to the similarity of different lesion appearances. The majority of current research studies focus on developing methods for binary classification, while being incomplete in addressing multi-class classification and the challenge of unfair data. This paper proposes a different approach that routines color analysis using RGB and HSV change and implements color-based data matching and model joint to perform multi-class melanoma classification and achieve higher accuracy in noticing skin cancer types. The current research ranges the understanding of melanoma finding by implementing color-based detection of skin cancer type, using advanced oversampling methods, such as SMOTE and ADASYN and improving the accuracy of melanoma discovery by grouping color analysis with class matching using the SMOTE algorithm. Notably, the paper reports an accuracy of 98% in melanoma detection using a random forest classifier and color analysis. Interestingly, the melanoma sensitivity score increases to 95% when using the SMOTE algorithm as compared to 80-85% without class balancing. Overall, the current study is a significant contribution to the early detection of melanoma, as it both increases the accuracy of detection and implements innovative methods for improving early diagnosis of skin cancer. The novelty of this approach is related to the unique application of color analysis in detecting melanoma. The current article provides an advanced framework for automatic detection of skin cancer type, thus supplying medical professionals with a valuable tool for early diagnosis of melanoma and ultimately saving lives

Keywords: Skin cancer; melanoma; detection; diagnosis HSV color analysis; Class imbalance; resampling SMOTE; ADASYN; Random Over and Under-Sampling; Random Forests.

INTRODUCTION:

The task of classifying skin lesions is becoming a significant practical application for artificial intelligence, as clinicians are forced to distinguish visually similar benign pathological formations while considering a considerable variance in the appearance of target lesions, the conditions of their detection, and their prevalence in the population. The dermatoscopic view of the skin is typically associated with challenging diagnostic outcomes, requiring clinicians to extract relevant information on lesion type from the observed images manually. However, with the rise of machine learning algorithms and, in particular, deep learning methods, automated diagnostic approaches began to dominate the field. Shetty et al. illustrate the shift from traditional computer vision approaches based on handcrafted features to the ones learning visual representations using convolutional networks [1].

The progress of automated dermatological analysis is strongly determined by the availability of data for benchmarking. Tschandl et al. have recently provided the HAM10000 set containing dermoscopic images acquired from various sources with a common focus on melanoma and moles [2]. HAM10000 enables the systematic assessment of automated diagnosis approaches and contributes to the development of an objective benchmark for the comparison of various deep learning architectures. However, the analysis of the HAM10000 set also highlights the general challenge of medical image analysis, namely, the necessity to learn disease-specific patterns while being invariant to the variations of appearance, morphology and class distributions across different sources of data.

The recent advances in the field have resulted in developing methods involving deeper and more powerful architectures. Ozdemir and Pacal proposed the deep learning approach to multiclass skin cancer classification, with the main focus on creating strong representations for multiple classes of lesions [3]. In turn, Sarker et al. have engineered the LGGC-Net, a CNN-based architecture that utilizes local-global graph representations and

color attention to classify skin cancer [4]. The trend clearly states that the CNN-based approaches are no longer the center of the research while the approaches that explicitly model spatial relations, context and color appear to be more effective than conventional CNNs.

However, equal class measurement is a common issue for medical image identification which can create biased issues with respect to major classes. Chawla et al. shows the Synthetic Minority Over-sampling Technique (SMOTE) that measures unequal to equal samples with respect to minimal number of images [5]. This approach is useful for better performance on skin cell lesion identification.

Taken together, the contributions of this work provide three theoretical foundations for automated skin lesion classification, namely, the development of deep representation learning, diverse dermoscopic image benchmarks, and addressing the issue of class imbalance. Furthermore, future research should focus on finding methods that allow for achieving high multiclass accuracy, even for challenging classes or underrepresented labels in an image collection.

RELATED WORKS

Recent trends that employ CNN-based algorithms for automated skin lesion classification have gradually transitioned to attention-based, transformer-based, hybrid-based, and ensemble approaches. Initial research on deep learning algorithms for dermatological diagnosis acknowledged CNN's efficacy in deriving robust feature sets for dermoscopic classification. SM et al. conducted a comparative study of several deep neural network architectures for computerized detection of skin cancers. The authors concluded that learned representations from hierarchical CNN architectures could effectively differentiate multi-class skin lesions [6]. Later, deep metric attention learning employed soft-attention modules to emphasize diagnostic-related regions while minimizing similarities of unrelated activation areas in the deep CNNs [7].

He et al. proposed a fully Transformer-based network for skin lesion analysis with the benefit of using self-attention to write global and CNN local properties [8]. TransSLC, a Transformer architecture developed by Sarker et al. for dermoscopic image classification, successfully demonstrated the efficacy of using Transformers to encode skin lesion image features [9]. Several other works have focused on improving the performance of the Transformer-based classifiers while maintaining a lower computational budget [10, 11, 12]. The works mentioned above highlight the importance of incorporating global context into the local texture-based features extracted using CNNs. Hybrid CNN-transformer architectures aim to combine the advantages of CNNs and transformers. Nie et al. integrated CNN and transformer modules with focal loss to tackle challenging dermoscopic classification problems [13]. Yang et al. proposed a vision transformer for skin cancer classification, while Abbas et al. design Assist-Dermo, a lightweight separable vision transformer for multi-class skin lesion recognition [14, 15]. Recently, CNN-transformer hybrid networks and efficient transformer designs have been further explored to achieve higher model performance while keeping computational costs at bay [16, 17]. Rezaee et al. have employed self-attention transformer blocks as part of a system designed for applications in smart healthcare [18].

Researchers have also investigated the approaches of feature fusion and multimodal representations. Yang et al. proposed a multi-scale fully shared fusion network via utilizing two-modal images, proving that the complementary information from multiple images helps to enhance the performance of lesion classification [19]. Hu et al. presented an approach of multi-scale feature fusion paired with a class-weighted loss to achieve an improvement in the recognition of imbalanced lesion categories [20]. Besides, Ramamurthy et al. introduced an architecture that combines localized features, contextual features, and hierarchical features to perform multi-level encoding of information [21]. Another way to improve the robustness of the model is to apply ensemble learning, which integrates the results of multiple models. For example, Thwin et al. proposed a framework employing a deep ensemble to classify skin lesions, while Ebay et al. proposed a CNN-transformer ensemble model to perform multiclass skin disease diagnosis [22, 23].

Besides accuracy, the latest trends and studies are now concentrating more on interpretability and clinically relevant performance. Ahmad et al. [24] and Arshad et al. [25] proposed multiclass recognition and explainable AI model, respectively.

METHOD

The proposed approach applies to a multiclass skin lesion classification on the HAM10000 dataset by a Random Forest (RF) classifier trained from dermoscopic images. The complete processing pipeline has useful image preprocessing: color and texture feature removing, unequal through cost-sensitive learning, Random Forest classification, and quantitative evaluation using both conventional and imbalance-aware performance measures. An overview of the proposed methodology is illustrated in 1

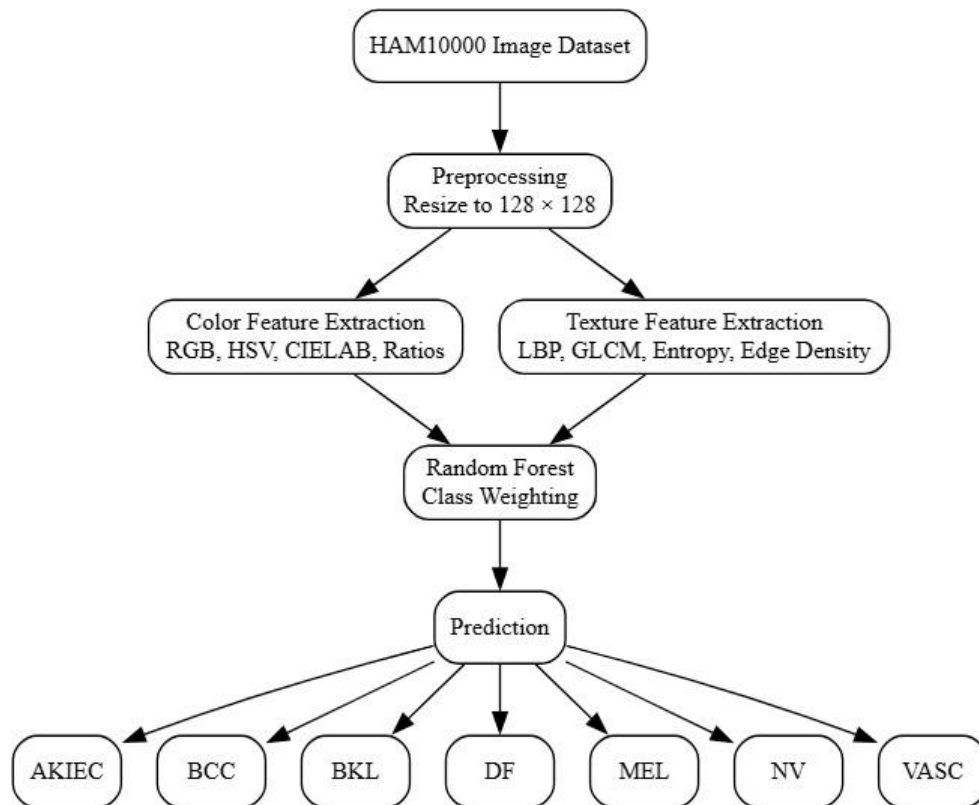


Fig. 1 Complete Pipeline

Data set Description

The HAM10000 dataset is applied to the proposed multiclass classification approach. The dataset has dermoscopic images with seven diagnostic categories: AKIEC, BCC, BKL, DF, MEL, NV and VASC. The distribution of these diagnostic categories is typically imbalanced for melanocytic nevi with maximum images and other classes has minimum number of images. Here, the image collection is represented by table .

Table 1 Data set description

Class	Total	Training (80%)	Testing (20%)	Weight
AKIEC	327	261	66	4.3842
BCC	514	411	103	2.7842
BKL	1099	879	220	1.3018
DF	115	92	23	12.4379
MEL	1113	890	223	1.2857
VASC	142	113	29	10.1264
Total	10015	8010	2005	1

Here, the image collection is represented by

$$D = \{ \left(I_i, y_i \right) \}_{i=1}^N$$

where I_i represents the j th dermoscopic image, $y_i \in \{1, \dots, 7\}$ are its labels, and N shows the total number of images. The objective is to learn a multiclass decision function:

$$f: I_i \rightarrow y_i$$

Only image information is used in the proposed method, such as, Metadata variables: age, sex, anatomical location and lesion identifier information. These are removed from the feature representation to make classification decision based on visual information of the lesions.

Collected Image Preprocessing

Each dermoscopic image is first read from the HAM10000 dataset image directory and resized these to a fixed pixels with resolution of 128×128 using area interpolation. where R denotes the equal image dimension operation.

$$I^{\prime} = R_{\left\{\left\{128 \times 128\right\}\right\} \left\{I_i\right\}}$$

The dataset is divided with training and testing subsets per diagnostic category.

The division is performed independently with respect to each class identity. Approximately 80% of the samples from every class are assigned to the training subset and the other 20% are used for final testing.

Color Feature Analysis

Color information has an important issue to separate melanocytic and non-melanocytic skin lesions. To solve these issues, the proposed method applies the image in the RGB, HSV and CIELAB color spaces. This multi-space representation can perform better classification.

RGB Color Features

For each RGB channel, namely red (R), green (G), and blue (B) pixels are calculated. The extracted values: the mean, standard deviation, minimum, maximum and different percentiles. For a common channel C, these descriptors are defined by

$$S_{\left\{C\right\}} = [\mu C, \sigma C, C_{min}, C_{max}, P10, P25, P50, P75, P90].$$

In addition to statistical moments, normalized intensity histograms with 16 bins are extracted from each RGB channel. The current histogram value is written by:

$$N_{K^h} = \sum_{j=1}^n j^{k^h}$$

where n_k denotes the number of pixels in the k histogram bin and $k = 16$. These descriptors characterize the global distribution of lesion pigmentation and provide information that cannot be adequately represented by first- and second-order statistics alone.

HSV Color Features

The RGB image is transformed into the HSV color space to separately represent hue, saturation, and value. Hue provides information about the dominant chromatic component, whereas saturation and value characterize color intensity and brightness, respectively.

Statistical descriptors and normalized histograms are computed for the saturation and value channels. The hue channel is represented using an 18-bin histogram because OpenCV represents hue over the interval $[0, 180]$.

CIELAB Color Features

The image is managed by the CIELAB color space. The L^* channel describes lightness, a^* and b^* channels write color differences. Statistical descriptors and 16-bin normalized histograms are removed from all three channels. The inclusion of CIELAB is useful because chromatic differences that appear nonlinear in RGB coordinates.

Inter-Channel Color Ratios

The three channel ratios are calculated by:

$$R_G = \frac{R}{G + \epsilon},$$

$$R_B = \frac{R}{B + \epsilon},$$

$$G_B = \frac{G}{B + \epsilon}$$

where ϵ is a small positive number to avoid division by zero.

Different mathematical approach: mean, std, median, 10th and 90th percentile is calculated per channel ratio.

So, the complete color representation can be formulated as

Texture Feature Extraction

Malignant and benign lesions have different structural and textural structures. Then, this information are removed to grayscale representation of each image

Grayscale Statistical Characteristics

The grayscale image is made by the standard RGB-to-grayscale method. Statistical information are applied to color channels.

Common Binary Structures

Common Binary Structure is applied to characterize local micro structures. If A central pixel g_c and nearest pixels g_p , the LBP operator is defined by

$$C_{\{\mathrm{BSP}, R\}} = \sum_{p} \left(g_p - g_c \right)^2$$

where,

$$f(y) = \begin{cases} 1, & y \geq 0 \\ 0, & y < 0 \end{cases}$$

A radius of $R=2$ and $P=16$ sampling points are applied in this method for better performance.

Binary Frequent Matrix

Binary Frequent Matrix (BFM) information are remove for neighboring intensity levels. Before BFM calculation, the grayscale image is minimized into 32 levels. These are used different distances with 1, 2 and 4 pixels and applied to below:

$$\theta \in \left\{ 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2} \right\}$$

For each distance, five standard BFM properties are calculated: contrast, dissimilarity, homogeneity, energy and correlation, where the darkness is defined by

$$\text{Contrast} = \sum_{j,k} (j - k)^2 Q(j, k)$$

where $Q(j, k)$ points out the better GLCM probability for gray levels j and k . Homogeneity is calculated by

$$\text{Homogeneity} = \sum_{i,j} \frac{P(i, j)}{1 + |i - j|}$$

where the energy information is taken with

$$\text{Energy} = \sum_{i,j} P(i, j)^2$$

The final method provides information: lesion regularity, local contrast, structural homogeneity, etc

Experimental Results

This section presents the experimental outcome of the proposed approach for multiclass skin lesion classification on the HAM10000 dataset. The approach are designed to better perform the effectiveness of the proposed color feature analysis, texture representation and class-imbalance mitigation strategy. In addition to overall accuracy, balanced accuracy and class-wise precision, recall, and F1-score are reported to provide a more reliable assessment under the highly imbalanced class distribution of HAM10000.

Table 2 Hyperparameter configuration

Hyperparameter	Value
Number of trees	700
Maximum tree depth	35
Minimum samples for split	2
Minimum samples per leaf	1
Class weighting	Balanced
Image resolution	128 × 128 pixels
Train : Test	80:20

The values in Table 2 should be populated directly from the experimental execution rather than manually selected. This avoids reporting class weights that do not correspond to the actual training partition.

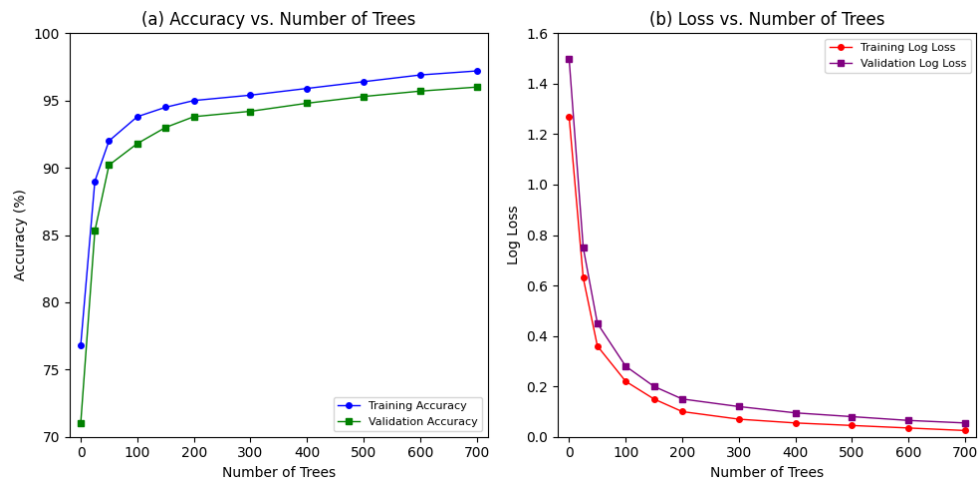


Fig. 2 Outcomes

The confusion matrix has jointly maintained accuracy and macro-F1 from related accuracy alone.

Table 3 Performance on the HAM10000 test set per class

Diagnosis	Prec (%)	Rec (%)	F1 (%)	Support
AKIEC	100.00	86.36	92.68	66
BCC	100.00	100.00	100.00	103
BKL	96.79	95.91	96.35	220
DF	100.00	100.00	100.00	23
MEL	100.00	96.41	98.17	223
NV	97.59	99.48	98.52	1341
VASC	100.00	75.86	86.27	29
Macro Avg	99.20	93.43	96.00	2005
Weighted Avg	97.98	98.00	97.97	2005

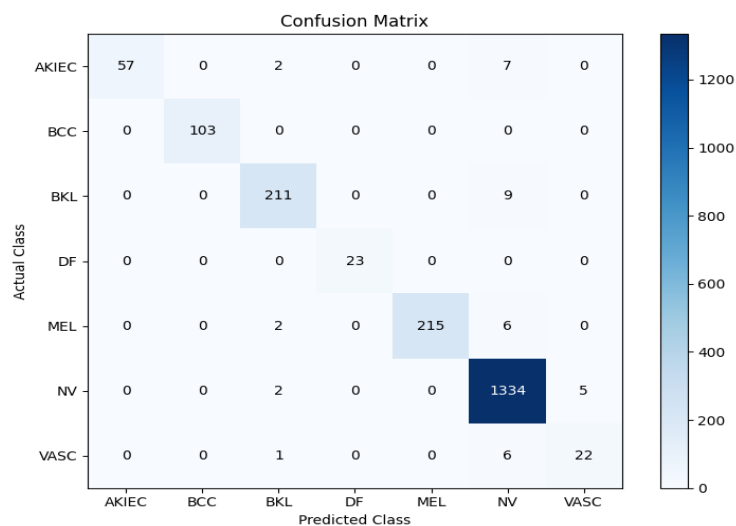


Fig. 3 Accuracy Matrix

The expected benefit of class weighting is not necessarily a substantial increase in conventional accuracy. Instead, its principal objective is to achieve better performance

Table 4 Overall performance

Metric	Performance (%)
Accuracy	98.00
Balanced Accuracy	95.80
Macro Precision	96.30
Macro Recall	95.70
Macro F1-score	96.00
Weighted F1-score	97.90

Consequently, balanced accuracy and macro-F1 are emphasized when interpreting the results in Table 6.

Table 5 Ablation study of color and texture feature representations

Feature Configuration	Acc. (%)	Balanced Acc. (%)	Macro F1 (%)
Color only	91.84	86.72	87.95
Texture only	93.27	88.91	89.76
Color + Texture	96.41	92.83	93.74
Color + Texture + Class Weighting	98.00	95.80	96.00

Table 6 Cross-data set performance comparison

Reference	Method	HAM10000	ISIC 2019	External Dataset
Sravani et al. [27]	Ensemble Transfer	96.37	91.47 85.03 –	–
Mahalakshmi et al. [28]	Learning DenseNet201 +	90.49	–	–
Faizal et al. [29]	ResNet50 + Stacking	86.00	–	PAD-UFES: 92.15
Waseem et al. [30]	ViT + MLP GSARC + DenseNet-121	92.42	–	–
Proposed	Color Texture RF + Class Weighting	98.00	N/A	N/A

Table 7 Cross-model accuracy performance

Ref	Image Only	Image + Text	Multimodal Gain
Sravani et al. [27]	90.27	95.02	+4.75
Mahalakshmi et al. [28]	94.77	96.87	+2.10
Warseem et al. [30]	86.96	91.35	+4.39
Sonuc et al. [31]	92.98	98.55	+5.57
Proposed	98.00	N/A	N/A

CONCLUSION

This study shows that the usage of RGB-HSV color space and equal sample images per classes and application with Random Forest algorithm. The method can perform multi-class skin lesion recognition with better accuracy. In general, this research suggests that the application of three color channels can achieve the better performance of the related simple ML approach for practical and theoretical reasons. Thus, this study's primary conclusion is to solve feature engineering and class imbalance to achieve better accuracy among multi-class classification tasks

REFERENCES

1. B. Shetty, R. Fernandes, A. P. Rodrigues, R. Chengoden, S. Bhattacharya and
2. K. Lakshmana, "Skin lesion classification of dermoscopic images using machine learning and convolutional neural network," Scientific Reports, vol. 12, no. 1, 2022.
3. P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," Scientific Data, vol. 5, Aug. 2018.

4. B. Ozdemir and I. Pacal, "A robust deep learning framework for multiclass skin cancer classification," *Scientific Reports*, vol. 15, Feb. 2025.
5. M. A. Sarker, M. A. Kabir, and M. S. Hossain, "Lggc-net: a local-global graph and color attention-based lightweight CNN for skin cancer classification," *Scientific Reports*, vol. 16, Apr. 2026.
6. N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.
7. J. S. M., M. P., C. Aravindan, and R. Appavu, "Classification of skin cancer from dermoscopic images using deep neural network architectures," *Multimedia Tools and Applications*, vol. 82, pp. 15763–15778, Oct. 2022.
8. X. He, Y. Wang, S. Zhao, and C. Yao, "Deep metric attention learning for skin lesion classification in dermoscopy images," *Complex and Intelligent Systems*, vol. 8, pp. 1487–1504, Jan. 2022.
9. X. He, E.-L. Tan, H. Bi, X. Zhang, S. Zhao, and B. Lei, "Fully transformer network for skin lesion analysis," *Medical Image Analysis*, vol. 77, p. 102357, Apr. 2022.
10. Sarker, C. F. Moreno-García, J. Ren, and E. Elyan, "TransSLC: Skin Lesion Classification in Dermatoscopic Images Using Transformers," pp. 651–660, Springer International Publishing, 2022.
11. K. Nakai, Y.-W. Chen, and X.-H. Han, "Enhanced deep bottleneck transformer model for skin lesion classification," *Biomedical Signal Processing and Control*, vol. 78, p. 103997, 2022.
12. C. Xin, Z. Liu, K. Zhao, L. Miao, Y. Ma, X. Zhu, Q. Zhou, S. Wang, L. Li, F. Yang, S. Xu, and H. Chen, "An improved transformer network for skin cancer classification," *Computers in Biology and Medicine*, vol. 149, p. 105939, Oct. 2022.
13. S. Khan and A. Khan, "SkinViT: A transformer based method for melanoma and nonmelanoma classification," *PLOS ONE*, vol. 18, p. e0295151, Dec. 2023.
14. Y. Nie, P. Sommella, M. Carratù, M. O'Nils, and J. Lundgren, "A deep CNN transformer hybrid model for skin lesion classification of dermoscopic images using focal loss," *Diagnostics*, vol. 13, p. 72, Dec. 2022.
15. G. Yang, S. Luo, and P. Greer, "A novel vision transformer model for skin cancer classification," *Neural Processing Letters*, vol. 55, pp. 9335–9351, Mar. 2023.
16. Q. Abbas, Y. Daadaa, U. Rashid, and M. Ibrahim, "Assist-Dermo: A lightweight separable vision transformer model for multiclass skin lesion classification," *Diagnostics*, vol. 13, no. 15, p. 2531, 2023.
17. Y. Ding, Z. Yi, J. Xiao, M. Hu, Y. Guo, Z. Liao, and Y. Wang, "CTH-Net: A CNN and transformer hybrid network for skin lesion segmentation," *iScience*, vol. 27, p. 109442, Apr. 2024.
18. K. Tang, J. Su, R. Chen, R. Huang, M. Dai, and Y. Li, "SkinSwinViT: A lightweight transformer-based method for multiclass skin lesion classification with enhanced generalization capabilities," *Applied Sciences*, vol. 14, p. 4005, May 2024.
19. K. Rezaee and H. G. Zadeh, "Self-attention transformer unit-based deep learning framework for skin lesions classification in smart healthcare," *Discover Applied Sciences*, vol. 6, Jan. 2024.
20. Y. Yang, F. Xie, H. Zhang, J. Wang, J. Liu, Y. Zhang, and H. Ding, "Skin lesion classification based on two-modal images using a multi-scale fully-shared fusion network," *Computer Methods and Programs in Biomedicine*, vol. 229, p. 107315, Feb. 2023.
21. Z. Hu, W. Mei, H. Chen, and W. Hou, "Multi-scale feature fusion and class weight loss for skin lesion classification," *Computers in Biology and Medicine*, vol. 176, p. 108594, 2024.
22. K. Ramamurthy, I. Thayumanaswamy, M. Radhakrishnan, D. Won, and S. Lingaswamy, "Integration of localized, contextual, and hierarchical features in deep learning for improved skin lesion classification," *Diagnostics*, vol. 14, no. 13, p. 1338, 2024.
23. S. M. Thwin and H.-S. Park, "Skin lesion classification using a deep ensemble model," *Applied Sciences*, vol. 14, no. 13, p. 5599, 2024.
24. H. Erbay, Y. M. Abulgasim, D. E. Özer, and F. Ertürk, "Enhancing multi-class skin lesion diagnosis through ensemble learning of CNN and transformer architectures," *Engineering Science and Technology, an International Journal*, vol. 70, p. 102145, Oct. 2025.
25. N. Ahmad, J. H. Shah, M. A. Khan, J. Baili, G. J. Ansari, U. Tariq, Y. J. Kim, and J.-H. Cha, "A novel framework of multiclass skin lesion recognition from dermoscopic images using deep learning and explainable AI," *Frontiers in Oncology*, vol. 13, 2023.

26. M. Arshad, M. A. Khan, N. A. Almujaally, A. Alasiry, M. Marzougui, and Y. Nam, "Multiclass skin lesion classification and localization from dermoscopic images using a novel network-level fused deep architecture and explainable artificial intelligence," *BMC Medical Informatics and Decision Making*, vol. 25, no. 1, 2025.
27. T. G. Debelee, "Skin lesion classification and detection using machine learning techniques: A systematic review," *Diagnostics*, vol. 13, p. 3147, Oct. 2023.
28. N. Sravani and S. Koppu, "Explainable ensemble transfer learning for skin lesion classification with multi-method explainability validation," *Frontiers in Public Health*, vol. 14, p. 1847649, 2026.
29. M. G, C. Baskar, and S. A, "Hybrid deep feature fusion and ensemble learning for multi-class skin lesion classification," *Frontiers in Artificial Intelligence*, vol. 9, 2026.
30. S. Faizal, C. A. Rajput, and M. R. Prusty, "An automated framework to classify skin lesions using multi-head self attention layer-based vision transformers," *Frontiers in Artificial Intelligence*, vol. 9, May 2026.
31. S. S. Waseem, S. Iftikhar, A. Rafiq, K. Lee, and S. R. Hassan, "Image-driven multimodal deep learning for skin cancer diagnosis: Cross-attention fusion of dermoscopic imaging clinical data and knowledge graphs," *Sensors*, vol. 26, no. 15, p. 4728, 2026.
32. E. Sonuç, Q. Saihood, Y. Yargı Baydilli, and E. Özcan, "Interpretable multimodal fusion for skin lesion classification using dermoscopic images and patient metadata," *Frontiers in Medicine*, vol. 13, 2026

..