

Research on Skin Mirror Classification Method Based on Feature Fusion and Machine Learning

Yixuan Xu, Xueri Li, Ruxia Wang, Lei Yang*

Department of Computer Science, Guangdong University of Science and Technology, Dongguan, China

**email address of corresponding author: lixueri@gdust.edu.cn*

Keywords: Skin Lesion Classification, Feature Fusion, Block Statistical Features, Color Histogram, Traditional Machine Learning

Abstract

In the field of computer vision, image feature extraction and representation remain the core factor determining classification performance. This study addresses misclassification issues caused by category similarity and feature overlap in dermatological lesion image classification by proposing a feature extraction method that integrates Block Statistical Features (BSF) with Color Histogram (CH). Using the Split Smol dataset containing nine types of skin diseases as the foundation, we calculated mean and variance through image segmentation, then constructed an 8-bin color histogram to build a fused feature vector. Comparative experiments were conducted on three traditional models-Extreme Trees (ET), AdaBoost (AB), and Histogram Gradient Boosting (HGB)-to evaluate classification performance between original pixels and fused features. Experimental results demonstrate that the fused features significantly improved overall accuracy, with F1-scores for easily confused categories like actinic keratosis and squamous cell carcinoma increasing by over 0.15 on average while markedly reducing false positives. The study reveals that complementary local texture and global color information can effectively enhance dermatological image differentiation without increasing model complexity, providing a robust solution for interpretable classification in low-sample scenarios.

1 Introduction

In recent years, the rapid development of artificial intelligence technology has profoundly transformed the paradigm of medical image analysis, particularly demonstrating tremendous application potential in dermatological lesion diagnosis [1]. Current research on skin lesion images both domestically and internationally primarily focuses on areas such as image segmentation [2], feature extraction [3], and diagnostic classification [4]. Researchers employ image segmentation techniques to accurately identify lesion regions and combine them with feature extraction algorithms to extract effective diagnostic features. Early studies predominantly concentrated on image segmentation and manual feature extraction. Hosny et al. [5] proposed applying Convolutional Neural Networks (CNN) AlexNet for image classification, utilizing multi-layer convolutions to automatically extract key features from lesion regions in skin images, achieving accurate lesion classification. Tabrizchi et al. [6] developed an improved network based on Visual Geometry Group Network (VGGNet), combining small convolutions with global average pooling layers to construct an efficient feature extraction network. The rich lesion features extracted through this approach significantly enhance classification accuracy. Tahir et al. [7] introduced the DSCC_Net model for skin lesion classification, which consists of convolutional blocks and dense blocks. Convolutional blocks extract local detail features from images through convolution operations, while dense blocks propagate feature maps generated at each layer to subsequent layers during forward propagation, enabling feature reuse and effective gradient transmission. This approach demonstrates

outstanding performance in skin lesion classification. The deep convolutional ensemble model DCENSnet proposed by Chanda et al. [8] integrates multiple deep convolutional neural networks (DCNNs), leveraging the diversity and complementarity between different networks to facilitate effective learning across various feature levels. Additionally, the traditional convolutional layers incorporated into DCENSnet enhance the model's ability to accurately extract critical features from skin lesion images, thereby improving classification accuracy. Benyahia et al. [9] conducted a detailed performance comparison of 17 pre-trained CNN models, identifying the most suitable feature extraction network for specific skin lesion datasets and demonstrating the critical role of feature extraction modules in classification performance. Although deep learning methods excel at automatic feature extraction, their "black box" nature leads to poor model interpretability and susceptibility to overfitting in small-sample scenarios. Consequently, some researchers have turned back to traditional machine learning models to explore diagnostic solutions with strong interpretability and lower training costs. Batista et al. [10] introduced evaluation criteria such as uncertainty, diversity, and representativeness to select samples with maximum information for training, significantly reducing the required annotated data volume. Experimental results demonstrate that AL achieves approximately 93% classification accuracy with minimal annotated samples and training iterations, effectively reducing expert annotation costs and providing an effective solution for early diagnosis of skin lesions. Khalid M. Hosny et al. [11] fine-tuned the last three layers of AlexNet and conducted data augmentation on the ISIC2018 dataset containing seven types of skin lesion images before

performing classification. However, existing studies generally suffer from limited feature representation capabilities, with traditional methods predominantly relying on raw pixels or single-type features (e.g., color, texture), making it difficult to distinguish highly similar lesion types (e.g., actinic keratosis and squamous cell carcinoma). Although deep models exhibit superior performance, their high demands for data volume and computational resources hinder deployment in primary healthcare settings. To address these challenges, this paper proposes a traditional machine learning method that integrates block statistical features with color histograms. Using the Split_smol nine-category dermatoscopy dataset as the basis, we systematically evaluate its classification performance across three models: Extra Trees, AdaBoost, and HistGradientBoosting. Experimental results show that our approach significantly improves overall accuracy without increasing model complexity, particularly excelling in distinguishing confounding categories like actinic keratosis and squamous cell carcinoma. This validates the effectiveness of integrating local texture with global color information. This study provides a new idea of lightweight, robust and clinically friendly interpretable classification of dermatoscopy images, which fills the gap in feature fusion mechanism of traditional machine learning.

2 Data

2.1 Experimental Data

Split Smol is a dermatology image dataset from Kaggle, widely used in research on skin disease classification and detection. The dataset contains 900 images covering nine categories of skin conditions: actinic keratosis (Actinic keratosis), atopic dermatitis (Atopic Dermatitis), benign keratosis (Benign Keratosis), dermatofibroma (Dermatofibroma), melanocytic nevus (Melanocytic nevus), melanoma (Melanoma), squamous cell carcinoma (Squamous cell carcinoma), tinea/ringworm/candidiasis (Tinea/Ringworm/Candidiasis), and vascular lesions (Vascular lesion). The Split smol dataset is divided into training and test sets in an 8:2 ratio. Specifically, the Actinic Keratosis training set contains 80 images with 20 test images, Atopic Dermatitis training set (81 images with 21 test images), Benign Keratosis training set (80 images with 20 test images), Dermatofibroma training set (80 images with 20 test images), Melanocytic Nevus training set (80 images with 20 test images), Melanoma training set (80 images with 20 test images), Squamous Cell Carcinoma training set (80 images with 20 test images), Tinea/Ringworm/Candidiasis training set (56 images with 20 test images), and Vascular Lesion training set (80 images with 20 test images). The total number of images in the training set is 697, while the test set has 181 images. Table 1 details the specific classifications of the Split_smol dataset.

Table 1 Specific classification of Split_smol dataset

Category of skin disease	training set	test set	total
Actinic keratosis	80	20	100
Atopic Dermatitis	81	21	102

Benign keratosis	80	20	100
Dermatofibroma	80	20	100
Melanocytic nevus	80	20	100
Melanoma	80	20	100
Squamous cell carcinoma	80	20	100
Tinea/Ringworm/Candidiasis	56	20	76
Vascular lesion	80	20	100
total	697	181	900

2.2 Data Preprocessing

In this study, all image data were uniformly processed as 256x256 pixel RGB format. This standardization ensures consistent spatial dimensions and channel information across samples, facilitating subsequent feature extraction and model training while avoiding computational complexity and performance fluctuations caused by image size variations. Subsequently, each image was divided into 16x16 blocks, with pixel mean and variance extracted as statistical features. The image is represented as $I \in R^{H \times W \times C}$, with a block size of $B \times B$. The mean $\mu_{i,j}^{(c)}$ and variance $\sigma_{i,j}^{2(c)}$ for the (i, j) block in the c -th channel can be expressed as formulas (1) and (2):

$$\mu_{i,j}^{(c)} = \frac{1}{B^2} \sum_{x=1}^B \sum_{y=1}^B I_{(iB+x), (jB+y), c} \quad (1)$$

$$\sigma_{i,j}^{2(c)} = \frac{1}{B^2} \sum_{x=1}^B \sum_{y=1}^B (I_{(iB+x), (jB+y), c} - \mu_{i,j}^{(c)})^2 \quad (2)$$

Here, H represents the height, W denotes the width, C stands for the number of input channels, i and j are block indices indicating their positions in the grid. C is the channel index, while x and y serve as relative pixel coordinate indices within blocks, used to compute sums, averages, and variances of all pixels. By concatenating the averages and variances of all blocks, we obtain the F_{block} statistical feature vector for the entire image.

$$F_{block} = [\mu_{1,1}^{(1)}, \sigma_{1,1}^{2(1)}, \dots, \mu_{H/B, W/B}^{(C)}, \sigma_{H/B, W/B}^{2(C)}] \quad (3)$$

Furthermore, to better capture the color distribution characteristics of dermatological images, we calculated the color histogram. Given that the image's pixel values range from $[0, L-1]$, they are divided into K equally wide intervals. In this study, K is set to 8. The normalized frequency of the c -th channel within the k -th bin is expressed as Equation (4):

$$h_{c,k} = \frac{1}{H \cdot W} \sum_{u=1}^H \sum_{v=1}^W g(I_{u,v,c} \in I_k) \quad (4)$$

Where $g(\cdot)$ is the indicator function, The formula for calculating the color histogram feature F_{hist} is shown in Equation (5):

$$F_{hist} = [h_{1,1}, \dots, h_{1,K}, h_{2,1}, \dots, h_{2,K}, \dots, h_{C,K}] \quad (5)$$

The final comprehensive feature vector is generated by fusing statistical block features with color histogram features of image data. This fusion approach significantly enhances both the comprehensiveness and discriminative power of dermatological image representation. Block statistical features capture local brightness distribution and texture variations in images, while color histogram features reflect overall color distribution trends. These complementary information sources enable the feature representation to encompass both detailed texture details and global color

characteristics. This fusion method not only improves discrimination between different skin disease categories but also enhances robustness against external factors such as lighting conditions, camera angles, and noise. Consequently, it effectively reduces model overfitting risks and provides more stable, adaptable input feature vectors for subsequent classification models.

Since labels in raw image data typically exist as strings or categorical names that cannot be directly recognized and processed by machine learning models, they must first be converted into integer indices using a LabelEncoder to enable numerical representation of different categories. Furthermore, to better compute metrics such as confusion matrices, classification reports, and multi-class ROC curves in multi-category classification tasks, one-hot encoding [12] is applied to the integer labels. This process maps each category to an independent binary vector, ensuring categorical independence and uniform representation space. Not only does this eliminate implicit numerical relationships between labels, but it also provides objective target representations for subsequent classifiers that meet multi-class classification requirements, facilitating both model performance evaluation and result interpretability.

3 Experiment

In this experiment, three traditional machine learning models-Extra Trees [13], AdaBoost [14], and HistGradientBoosting [15] -were employed for dermatological image classification. The experimental environment consisted of Ubuntu 20.04.6 LTS operating system, Python 3.9, scikit-learn 1.3.2, an i7-12700H CPU with 32 GB of RAM, and a fixed global random seed of 129. Specifically, the Extra Trees model was configured with 200 decision trees using `max_features="sqrt"` and multi-threading enabled. AdaBoost utilized decision trees with `max_depth=3` as base learners, performed 100 iterations, and had a learning rate of 0.8. HistGradientBoosting was set to `max_iter=100`, `learning_rate=0.1`, `max_depth=3`, `min_samples_leaf=20`, and `l2_regularization=1.0`. All models used flattened original pixel images as input features and were trained and evaluated on identical training and test datasets. The confusion matrix generated by the Extra Trees model is illustrated in Figure 1.

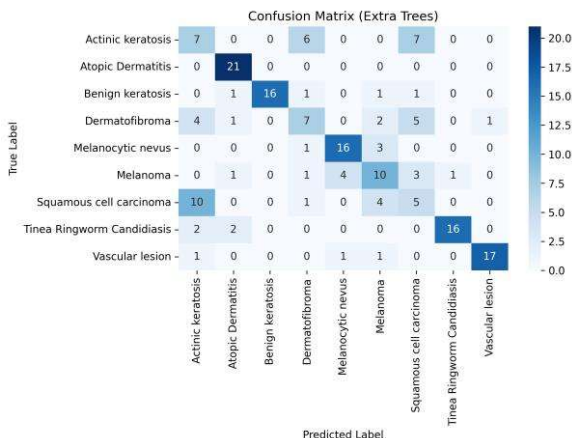


Figure 1 Confusion matrix trained with Extra Trees

As shown in Figure 1, the model demonstrates optimal classification performance for Atopic Dermatitis and Vascular lesions with perfect prediction accuracy, indicating distinct and easily distinguishable characteristics of these categories. However, significant confusion occurs in predicting Actinic Keratosis, Dermatofibroma, and Squamous cell carcinoma. For instance, Actinic Keratosis is frequently misclassified as Dermatofibroma or Squamous cell carcinoma, while some Dermatofibroma cases are mistakenly labeled as Actinic Keratosis, Melanoma, or Squamous cell carcinoma. A substantial number of Squamous cell carcinoma samples are incorrectly categorized as Actinic Keratosis. This misclassification primarily stems from overlapping color patterns, similar textures, and morphological features across these categories, combined with potential imbalanced training datasets. Additionally, the experiment employs raw pixel expansion as a feature representation without enhancing local texture and color distributions, which hinders the model's ability to differentiate similar classes. Furthermore, lighting conditions, occlusions, or background interference in images may increase misclassification risks. The ROC curve trained using Extra Trees is illustrated in Figure 2.

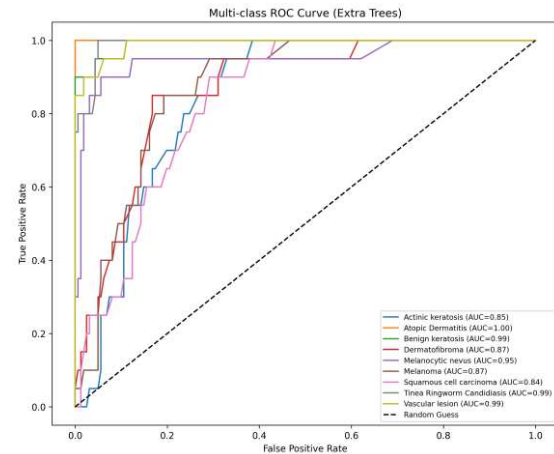


Figure 2 ROC curve trained with Extra Trees

As shown in Figure 2, the ROC curves for Atopic Dermatitis and Vascular lesions demonstrate higher AUC values, indicating these categories exhibit distinct features that are relatively easy to distinguish. In contrast, the ROC curves for Actinic Keratosis, Dermatofibroma, and Squamous cell carcinoma show steeper slopes with lower AUC values, reflecting challenges in model generalization and recognition difficulty. The confusion matrix generated through AdaBoost training is illustrated in Figure 3.

As shown in Figure 3's confusion matrix of the AdaBoost original pixel model, the model demonstrates optimal performance in Atopic Dermatitis classification with accurate prediction of all samples, indicating distinct and easily distinguishable original pixel features for this category. However, significant prediction confusions occur in Actinic Keratosis, Dermatofibroma, and Squamous cell carcinoma. For instance, Actinic Keratosis samples were misclassified as Dermatofibroma and Squamous cell carcinoma, while some Dermatofibroma samples were incorrectly labeled as Actinic

Keratosis, Melanoma, or Squamous cell carcinoma. Similarly, numerous Squamous cell carcinoma samples were misclassified as Actinic Keratosis, revealing substantial overlap in original pixel feature spaces among these categories. Other classes like Benign keratosis, Melanocytic nevus, Melanoma, Tinea/Ringworm/Candidiasis, and Vascular lesions showed moderate prediction accuracy but still experienced minor sample misclassifications into similar categories. In general, the effect of directly using original pixels to train AdaBoost model is limited, and some categories are difficult to distinguish, indicating that the original pixel features can not fully capture key information such as texture and color. The ROC curve obtained by AdaBoost training is shown in Figure 4.

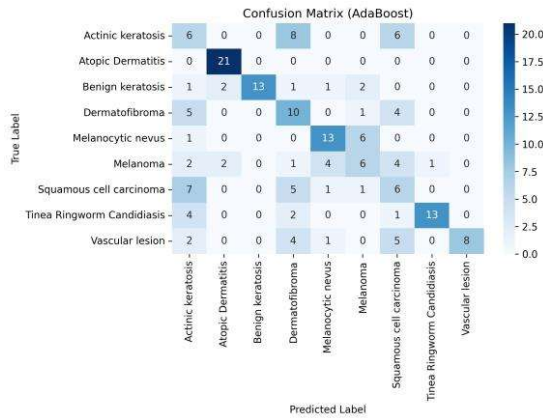


Figure 3 Confusion matrix obtained by AdaBoost training

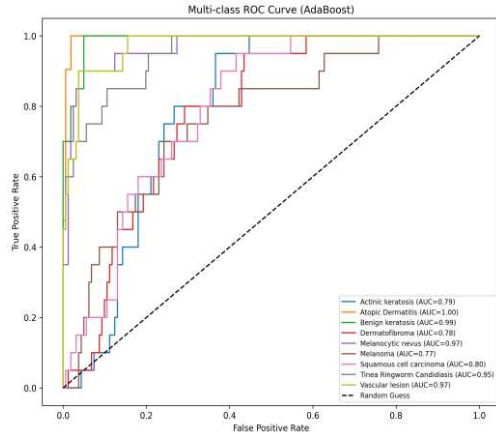


Figure 4 ROC curve obtained by AdaBoost training

The ROC curves of AdaBoost's original pixel model reveal distinct performance characteristics across categories. The Atopic Dermatitis category demonstrates the highest precision, with its ROC curve approaching the top-left corner and achieving an AUC value of 1, indicating highly accurate differentiation. In contrast, the Actinic Keratosis, Dermatofibroma, and Squamous cell carcinoma categories show relatively flat curves with significantly lower AUC values than others, suggesting weaker predictive capabilities and higher misclassification risks. Other categories including Benign keratosis, Melanocytic nevus, Melanoma,

Tinea/Ringworm/Candidiasis, and Vascular lesions exhibit moderate AUC levels, demonstrating some predictive distinguishability while still showing partial classification inaccuracies. The confusion matrix generated through HistGradientBoosting training is illustrated in Figure 5.

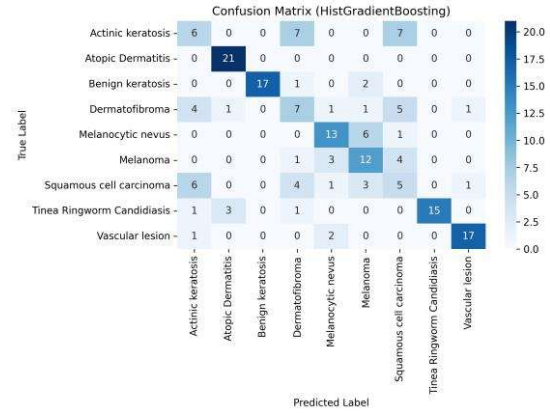


Figure 5 Confusion matrix obtained by HistGradientBoosting training

As shown in Figure 5, the HistGradientBoosting model demonstrates the best performance in classifying Atopic Dermatitis, with all samples accurately predicted, indicating distinct and easily distinguishable category features. However, significant confusion persists in predicting Actinic Keratosis, Dermatofibroma, and Squamous cell carcinoma. For instance, some Actinic Keratosis samples were misclassified as Dermatofibroma and Squamous cell carcinoma, while Dermatofibroma samples were incorrectly categorized as Actinic Keratosis, Squamous cell carcinoma, and Melanoma. Similarly, numerous Squamous cell carcinoma cases were mislabeled as Actinic Keratosis and Dermatofibroma, highlighting significant overlap in their feature spaces. Other categories like Benign keratosis, Melanocytic nevus, Melanoma, Tinea/Ringworm/Candidiasis, and Vascular lesions showed moderate prediction accuracy but still experienced minor misclassifications into similar categories. The ROC curve obtained by HistGradientBoosting training is shown in Figure 6.

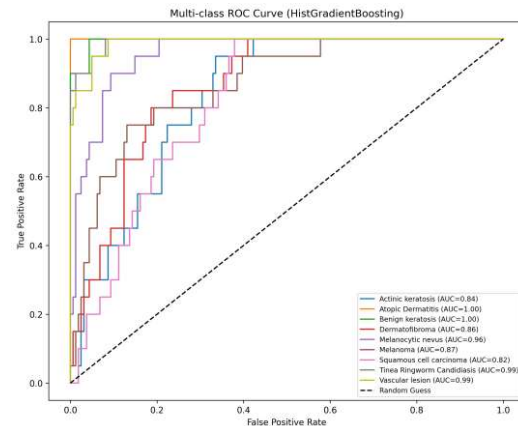


Figure 6 ROC curve trained with HistGradientBoosting

As shown in Figure 6, most classification curves are clearly above the diagonal line, indicating the model demonstrates strong overall classification capability. Notably, the AUC values for Atopic Dermatitis and Benign Keratosis reached 1, demonstrating excellent discriminative performance. However, for categories with high confusion potential such as Actinic Keratosis, Dermatofibroma, and Squamous Cell Carcinoma, the ROC curves showed lower performance compared to ideal models, with smaller AUC values. This aligns with the analysis from the confusion matrix, revealing that raw pixel features struggle to adequately represent local structural and textural information in images, resulting in suboptimal classification outcomes for some similar categories.

Experimental results from the original model indicate that while trained directly on expanded raw pixels, the model demonstrated strong performance in Atopic Dermatitis and Vascular Lesions but exhibited significant confusion for Actinic Keratosis, Dermatofibroma, and Squamous Cell Carcinoma categories, resulting in limited overall classification effectiveness. This suggests that raw pixel features fail to adequately capture local structural, textural, and chromatic distribution information in images. To address this, we subsequently employed a fusion approach combining block statistical features with color histogram features to enhance feature representation capabilities, thereby improving the model's discriminative performance in challenging category classifications. The fused feature matrix of the Extra Trees algorithm after this integration is illustrated in Figure 7.

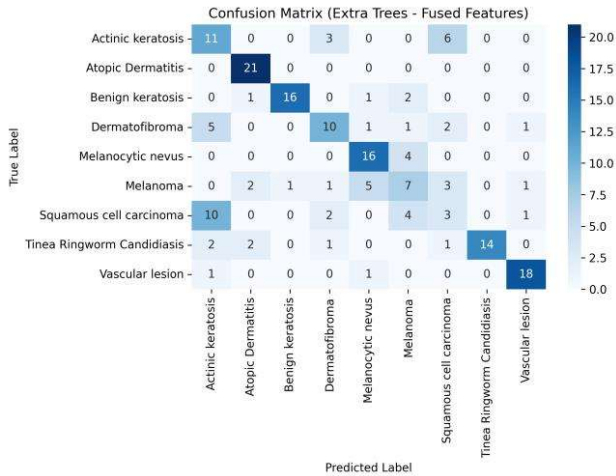


Figure 7 Confusion matrix of Extra Trees after integrating statistical features with color histogram features

As shown in Figure 7, the model achieved perfect classification accuracy for Atopic Dermatitis and Vascular lesions, demonstrating their distinct characteristics and ease of differentiation. The prediction accuracy also improved for categories such as Melanocytic Nevus, Melanoma, and Benign Keratosis. Specifically, Melanocytic Nevus maintained high accuracy rates, while Melanoma showed reduced misclassifications. Minor improvements were observed in Benign Keratosis misclassifications. Confusion between Actinic Keratosis, Dermatofibroma, and Squamous

Cell Carcinoma was alleviated, with fewer instances of Actinic Keratosis being misclassified as Dermatofibroma or Squamous Cell Carcinoma, and decreased mutual confusion between Dermatofibroma and Squamous Cell Carcinoma.

Compared to the original Extra Trees model, the improvement in fusion features stems from two key factors. Firstly, block statistical features better capture local texture information in images, while color histograms provide global color distribution characteristics. These complementary features enhance the model's ability to distinguish between different types of features. Secondly, fused features improve the expressive power of the feature space, enabling clearer differentiation between categories that were prone to overlap in the original block mean/standard deviation space. This reduces misclassifications between easily confused categories and ultimately improves overall classification accuracy. The ROC curve of the Extra Trees model after integrating block statistical features with color histogram features is illustrated in Figure 8.

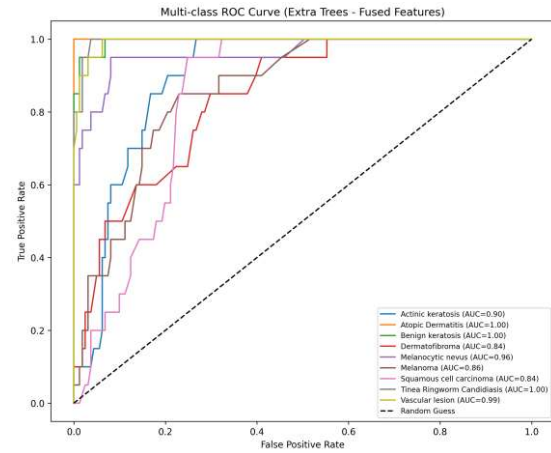


Figure 8 ROC curve of Extra Trees after integrating statistical features and color histogram features

The ROC curves of the Extra Trees model, which integrates statistical features with color histogram features, demonstrate significant improvements in overall performance across multi-category classifications. Compared to the original feature model, most categories show enhanced AUC values, particularly for previously confusing cases like Actinic Keratosis, Dermatofibroma, and Squamous Cell Carcinoma. These curves now cluster closer to the upper-left corner, indicating improved diagnostic accuracy. Atopic Dermatitis maintains its characteristic curve at 1, preserving its advantage of clear and easily distinguishable features. The overall ROC curve becomes smoother and shifts upward, suggesting higher true positive rates with lower false positives, reflecting notable performance enhancements. Figure 9 illustrates the confusion matrix of AdaBoost after applying the integrated statistical-color histogram feature approach.

The AdaBoost confusion matrix comparing fusion features with original features demonstrates that feature fusion significantly improves classification accuracy. While the prediction of Atopic Dermatitis remains entirely accurate, the

number of misclassifications for Actinic Keratosis, Dermatofibroma, and Squamous cell carcinoma-originally prone to confusion-has decreased. The classification performance for Melanocytic nevus, Melanoma, and Tinea/Ringworm/Candidiasis also shows slight improvements or remains stable. Figure 10 illustrates the ROC curves of AdaBoost after applying the method combining fused block statistical features with color histogram features.

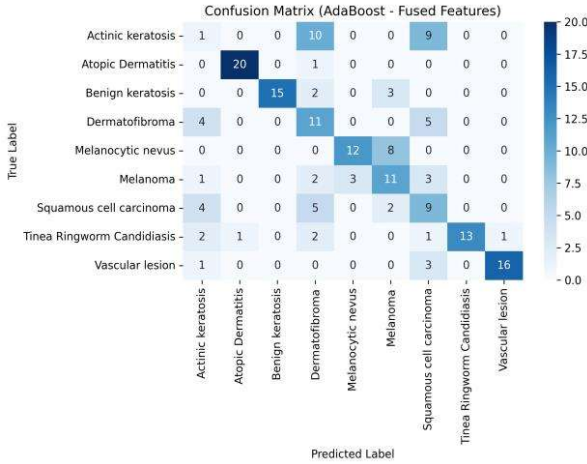


Figure 9 AdaBoost confusion matrix after integrating statistical features and color histogram features

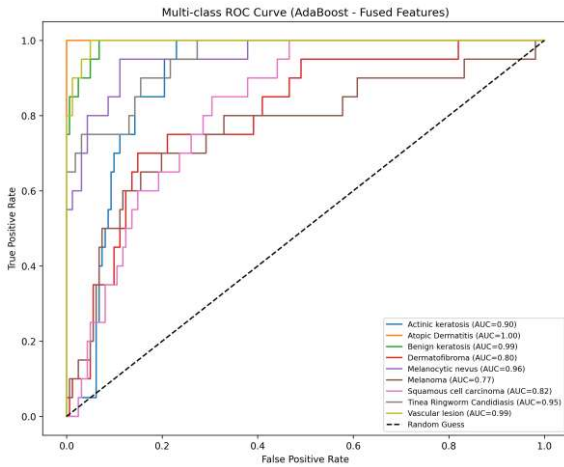


Figure 10 ROC curve of AdaBoost after integrating statistical features and color histogram features

The AdaBoost ROC curves after integrating statistical features with color histograms demonstrate improved overall AUC values across categories. Notably, for previously confusing categories like Actinic Keratosis, Dermatofibroma, and Squamous Cell Carcinoma, the curves show closer alignment to the upper-left corner, indicating enhanced discriminative power with reduced false positives and higher true positives. In contrast, AdaBoost ROC curves without integrated features exhibit significant deviations from the upper-left corner in some categories, showing lower AUC values that reflect suboptimal performance in distinguishing similar classes. Overall, the steeper and broader ROC curves after feature integration demonstrate stronger discriminative

capabilities and more reliable predictions across multiple categories. The confusion matrix of HistGradientBoosting using integrated statistical and color histogram features is illustrated in Figure 11.

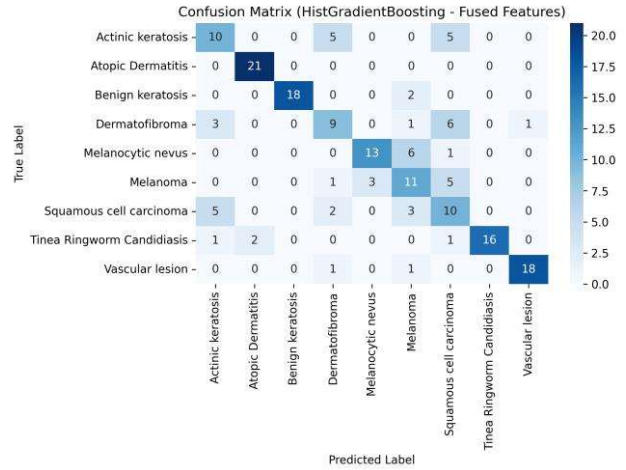


Figure 11 HistGradientBoosting confusion matrix using the method of combining block statistical features and color histogram features

By comparing the HistGradientBoosting confusion matrix of original features with integrated block statistical features combined with color histogram features, the performance improvement of fusion features becomes evident. Taking specific categories as examples: Correct classifications for Actinic Keratosis increased from 6 to 10, while misclassifications into Dermatofibroma and Squamous cell carcinoma were significantly reduced. Correct classifications for Benign Keratosis rose from 17 to 18 with fewer false positives. Dermatofibroma saw an increase from 7 to 9, and Squamous cell carcinoma improved from 5 to 10. Meanwhile, Atopic Dermatitis maintained perfect classification accuracy, while Vascular lesions achieved a 15-to-16 improvement in correct classifications. Overall, the integration of features enhanced correct classifications across most categories and minimized misclassifications between major confounding groups. This improvement stems from the complementary nature of block statistical features capturing local texture and brightness variations, and color histograms providing global chromatic distribution information. The synergy between these approaches strengthens the model's discriminative power when distinguishing similar categories, particularly effective for common confounders like Actinic Keratosis and Squamous cell carcinoma. Consequently, this approach comprehensively improves multi-category classification accuracy and robustness. The ROC curve of HistGradientBoosting after using the fusion block to count features and color histogram features is shown in Figure 12. As shown in Figure 12, AUC values improved across nearly all categories, particularly for previously ambiguous types. For instance, Actinic keratosis's AUC rose from 0.82 to 0.91, while squamous cell carcinoma's AUC increased from 0.78 to 0.88. Distinctive categories like Atopic Dermatitis and Vascular lesions maintained AUC values of 1 after applying fusion features. Compared with ROC curves without fusion

features, the lower-tier categories under original features exhibited flatter curves deviating from the upper-left quadrant, with lower AUC values, indicating higher false positives and false negatives in distinguishing these categories. By fusing statistical features with color histograms, the model integrates local texture, brightness distribution, and global color information, significantly enhancing its ability to differentiate similar or confusing categories.

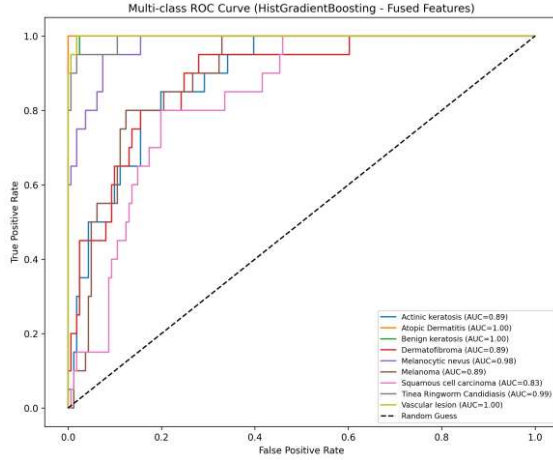


Figure 12 ROC curve of HistGradientBoosting after integrating statistical features and color histogram features

In comprehensive analysis of dermatological dataset studies, three models-Extra Trees, AdaBoost, and HistGradientBoosting-showed significant performance improvements when integrating block statistical features with color histogram features. The combined feature representation provides more comprehensive image information, encompassing both local texture details and global color distributions. This enhanced capability enables better differentiation between similar categories such as Actinic Keratosis, Squamous Cell Carcinoma, and Dermatofibroma, with prediction accuracy markedly improved. Specifically, Extra Trees achieved a notable increase in accurate predictions for Actinic Keratosis from 7 to 11. ROC curves demonstrate enhanced AUC values across nearly all categories, accompanied by reduced false positives and increased true positives, indicating strengthened predictive confidence and stability. These improvements primarily stem from the integrated features' ability to improve category discrimination, thereby enhancing classification accuracy and prediction reliability. Table 2 presents the prediction accuracy of the three original feature models, while Table 3 shows the performance after incorporating block statistical and color histogram features.

Table 2 Predictive accuracy of three models using original data features

models	classification	precision	recall	f1-sc support ore
Extra Trees	Actinic keratosis	0.29	0.35	0.32 20
	Atopic	0.81	1.00	0.89 21
	Dermatitis			

AdaBoost	Benign keratosis	1.00	0.80	0.89 20
	Dermatofibroma	0.41	0.35	0.38 20
	Melanocytic nevus	0.76	0.80	0.78 20
	Melanoma	0.48	0.50	0.49 20
	Squamous cell carcinoma	0.24	0.25	0.24 20
	Tinea/Ringworm /Candidiasis	0.94	0.80	0.86 20
	Vascular lesion	0.94	0.89	0.89 20
	accuracy	/	/	0.64 181
	Actinic keratosis	0.21	0.30	0.25 20
	Atopic	0.84	1.00	0.91 21
HistGradientBoosting	Dermatitis			
	Benign keratosis	1.00	0.65	0.79 20
	Dermatofibroma	0.32	0.50	0.39 20
	Melanocytic nevus	0.65	0.65	0.65 20
	Melanoma	0.38	0.30	0.33 20
	Squamous cell carcinoma	0.23	0.30	0.26 20
	Tinea/Ringworm /Candidiasis	0.93	0.65	0.76 20
	Vascular lesion	1.00	0.40	0.57 20
	accuracy	/	/	0.53 181
	Actinic keratosis	0.33	0.30	0.32 20
	Atopic	0.84	1.00	0.91 21
	Dermatitis			
	Benign keratosis	1.00	0.85	0.92 20
	Dermatofibroma	0.33	0.35	0.34 20
	Melanocytic nevus	0.65	0.65	0.65 20
	Melanoma	0.50	0.60	0.55 20
	Squamous cell carcinoma	0.23	0.25	0.24 20
	Tinea/Ringworm /Candidiasis	1.00	0.75	0.86 20
	Vascular lesion	0.89	0.87	0.87 20
	accuracy	/	/	0.62 181

Table 3 Prediction accuracy of the three models after integrating block statistical features and color histogram features

models	classification	precision	recall	f1-sc support ore
Extra Trees	Actinic keratosis	0.38	0.55	0.45 20
	Atopic	0.81	1.00	0.89 21
	Dermatitis			
	Benign keratosis	0.94	0.80	0.86 20
	Dermatofibroma	0.59	0.50	0.54 20
	Melanocytic nevus	0.67	0.80	0.73 20
	Melanoma	0.39	0.35	0.37 20
	Squamous cell carcinoma	0.20	0.15	0.17 20
	Tinea/Ringworm /Candidiasis	1.00	0.70	0.82 20
	Vascular lesion	0.86	0.88	0.88 20
	accuracy	/	/	0.64 181

AdaBoost	Actinic keratosis	0.08	0.05	0.06	20
	Atopic	0.95	0.95	0.95	21
	Dermatitis				
	Benign keratosis	1.00	0.75	0.86	20
	Dermatofibroma	0.33	0.55	0.42	20
	Melanocytic	0.80	0.60	0.69	20
	nevus				
	Melanoma	0.46	0.55	0.50	20
	Squamous cell	0.30	0.45	0.36	20
	carcinoma				
HistGradientBoosting	Tinea/Ringworm	1.00	0.65	0.79	20
	/Candidiasis				
	Vascular lesion	0.80	0.86	0.86	20
	accuracy	/	/	0.60	181
	Actinic keratosis	0.53	0.50	0.51	20
	Atopic	0.91	1.00	0.95	21
	Dermatitis				
	Benign keratosis	1.00	0.90	0.95	20
	Dermatofibroma	0.50	0.45	0.47	20
	Melanocytic	0.81	0.65	0.72	20
	nevus				
	Melanoma	0.46	0.55	0.50	20
	Squamous cell	0.36	0.50	0.42	20
	carcinoma				
	Tinea/Ringworm	1.00	0.80	0.89	20
	/Candidiasis				
	Vascular lesion	0.90	0.92	0.92	20
	accuracy	/	/	0.70	181

As shown in Table 2 and Table 3, all three models demonstrated improved overall performance after incorporating both fused statistical features and color histogram features. Specifically, Extra Trees showed enhanced performance in hard-to-distinguish categories: the F1-score for Actinic Keratosis increased from 0.32 to 0.45, while Dermatofibroma's F1-score rose from 0.38 to 0.54. AdaBoost achieved an overall accuracy improvement from 0.53 to 0.60, with Atopic Dermatitis showing a significant F1-score increase from 0.91 to 0.95, Dermatofibroma's F1-score rising from 0.39 to 0.42, and melanoma also achieving notable progress. HistGradientBoosting saw its overall accuracy improve from 0.62 to 0.70, with consistent gains across most categories. For instance, Actinic Keratosis's F1-score jumped from 0.32 to 0.51, while Benign Keratosis and Tinea/Ringworm/Candidiasis further strengthened their performance, demonstrating balanced and robust improvements. The fundamental reason for this improvement lies in the fact that fused block statistical features effectively capture local texture and brightness details of images, while color histograms provide comprehensive color distribution information. The combination of these two types of features significantly enriches the feature representation space, enhancing the model's ability to distinguish categories with similar appearances and complex textures. Consequently, this approach improves overall classification accuracy and prediction reliability.

4 Summary

This study systematically investigates the performance

improvement of traditional machine learning models when incorporating fused features for automated classification of skin lesion images. Using the Split Smol dataset from Kaggle, which contains 900 dermatological images across 9 categories, we conducted experiments with three models: Extra Trees, AdaBoost, and HistGradientBoosting. Initial training with raw pixel flattening features showed good performance in some categories (e.g., atopic dermatitis, vascular lesions), but resulted in significant confusion between similar conditions like actinic keratosis, cutaneous fibroma, and squamous cell carcinoma. To enhance feature representation, we proposed a fusion method combining block statistical features (capturing local texture and brightness variations) with color histogram features (reflecting global color distribution). This complementary approach significantly improved model accuracy in distinguishing similar lesions. Experimental results demonstrated enhanced overall accuracy across all three models post-feature fusion. Notably, HistGradientBoosting achieved the most substantial improvements, with F1-scores rising from 0.32 and 0.24 to 0.51 and 0.42 respectively in confusing categories like actinic keratosis and squamous cell carcinoma, while its ROC curves approached ideal performance levels.

However, this study still has certain limitations. First, the relatively small dataset size and uneven distribution of categorical samples may affect the model's generalization ability. Second, while the proposed feature fusion method is effective, it has not been compared with deep learning's automatic feature extraction capabilities, indicating potential performance limits. Additionally, the experiments lacked data augmentation or cross-validation, resulting in insufficient evaluation of model stability. Future work will expand the dataset scale, introduce more balanced sample distributions, and integrate deep learning methods (such as CNN and Vision Transformer) for feature extraction and classification to further enhance model performance. Simultaneously, incorporating multimodal information (e.g., patient medical history and lesion location) and interpretability analysis will promote practical application and credibility improvement of the model in clinical environments.

In conclusion, this paper verifies the effectiveness of integrating local and global features in skin disease classification, shows that traditional machine learning models are still competitive with enhanced features, and provides an interpretable and computationally efficient scheme for subsequent research.

5 Acknowledgements

This work was supported by the General Project of Guangdong University of Science and Technology (Grant No.GKY-2023KYYBK-20,GKY-2024KYYBK-16,GKY-2024KYONK-25and GKY-2024KYYBK-29).

6 Reference

[1] Najjar R. Redefining radiology: a review of artificial intelligence integration in medical imaging. *Diagnostics*,

2023, 13(17): 2760.

[2] Rehan M A, Li K, Chen P. Multimodal Pathological Image Segmentation Using the Integration of Trans MMY Net and Patient Metadata. *Electronics*, 2025, 14(12): 2369-2369.

[3] Mjihad A, Muñoz R A. Automated Quality Control of Candle Jars via Anomaly Detection Using OCSVM and CNN-Based Feature Extraction. *Mathematics*, 2025, 13(15): 2507-2507.

[4] Annapoorani G, Manikandan P, Genitha H C. Evolving medical image classification: a three-tiered framework combining MSPLnet and IRNet-VGG19. *Evolving Systems*, 2024, 16(1): 19-19.

[5] Hosny K M, Kassem M A, Fouad M M. Classification of skin lesions into seven classes using transfer learning with AlexNet. *Journal of Digital Imaging*, 2020, 33: 1325-1334.

[6] Tabrizchi H, Parvizpour S, Razmara J. An improved VGG model for skin cancer detection. *Neural Processing Letters*, 2023, 55(4): 3715-3732.

[7] Tahir M, Naeem A, Malik H, et al. DSCC_Net: multi-classification deep learning models for diagnosing of skin cancer using dermoscopic images. *Cancers*, 2023, 15(7): 2179.

[8] Chanda D, Onim M S H, Nyeem H, et al. DCENSnet: a new deep convolutional ensemble network for skin cancer classification. *Biomedical Signal Processing and Control*, 2024, 89: 105757.

[9] Benyahia N, et al. Pretrained CNN models for skin lesion classification: A comparison study. *International Journal of Computer Vision*, 2020, 128(5): 925-937.

[10] Batista L G, Bugatti P H, Saito P T M. Classification of skin lesion through active learning strategies. *Computer Methods and Programs in Biomedicine*, 2022, 226: 107122.

[11] Hosny K M, Kassem M A, Fouad M M. Classification of skin lesions into seven classes using transfer learning with AlexNet. *Journal of Digital Imaging*, 2020, 33(5): 1325-1334.

[12] Jawad M, Ming L, Mudassar R. An accurate detection of tool wear type in drilling process by applying PCA and one-hot encoding to SSA-BLSTM model. *The International Journal of Advanced Manufacturing Technology*, 2021, 118(11-12): 3897-3916.

[13] Aziz T, Camana R M, Garcia E C. REM-Based Indoor Localization with an Extra-Trees Regressor. *Electronics*, 2023, 12(20).

[14] Attou H, Guezzaz A, Benkirane S. A New Secure Model for Cloud Environments Using RBFNN and AdaBoost. *SN Computer Science*, 2025, 6(2): 188-188.

[15] Theerthagiri P. Liver disease classification using histogram-based gradient boosting classification tree with feature selection algorithm. *Biomedical Signal Processing and Control*, 2025, 100(PC): 107102-107102.