Article

Hybrid Transform-Based Feature Extraction for Skin Lesion Classification Using RGB and Grayscale Analysis

Luis Felipe López-Ávila and Josué Álvarez-Borrego [*]

Article

Hybrid Transform-Based Feature Extraction for Skin Lesion Classification Using RGB and Grayscale Analysis

Luis Felipe López-Ávila and Josué Álvarez-Borrego *

Centro de Investigación Científica y de Educación Superior de Ensenada, B. C. (CICESE), Carretera Ensenada-Tijuana No. 3918, Zona Playitas, Ensenada, B. C. 22860 México

*   Correspondence: josue@cicese.mx; Tel.: +52-(646)-175-0500

Featured Application: A technique for developing a potential automated system for classifying skin lesions.

Abstract: Automated skin lesion classification using machine learning techniques is crucial for early and accurate skin cancer detection. This study proposes a hybrid method combining the Hermite, Radial Fourier–Mellin, and Hilbert transform to extract comprehensive features from skin lesion images. By separating the images into red, green, and blue (RGB) channels and grayscale, unique textural and structural information specific to each channel is analyzed. The Hermite transform captures localized spatial features, while the Radial Fourier–Mellin and Hilbert transforms ensure global invariance to scale, translation, and rotation. Texture information for each channel is also obtained based on the Local Binary Pattern (LBP) technique. We applied our hybrid transform-based feature extraction approach to multiple lesion classes using the International Skin Imaging Collaboration (ISIC) 2019 dataset, preprocessed with data augmentation. Experimental results demonstrate that the proposed method improves classification accuracy and robustness, highlighting its potential as a non-invasive AI-based tool for dermatological diagnosis.

Keywords: skin lesion classification; Hermite transform; radial Fourier-Mellin transform

1. Introduction

Early detection of skin cancer is crucial for improving patient outcomes, as it remains a common and potentially deadly disease worldwide. To support dermatologists in diagnosis, automated classification of skin lesions has gained significant attention through the integration of machine learning (ML) and advanced digital image processing. By leveraging mathematical transformations and feature extraction techniques, ML models have shown strong potential in identifying diagnostic patterns from digital image data, making them practical tools for classifying skin lesions [1–3].

Integral transforms, such as Fourier, Mellin, and Hilbert, have proven particularly useful in generating feature sets that are invariant to changes in rotation, scale, and translation—properties that are essential for analyzing skin lesions, which exhibit considerable variability in shape, size, and orientation [4–6]. Frequency-based patterns, essential for differentiating various lesion textures and structures, have been analyzed using the Fourier transform, which is widely used in image processing [7,8]. The Mellin transform has also been applied to address the issue of varying lesion sizes in images, introducing scale invariance to the extracted features and enhancing their robustness [9–11]. Hilbert transform complements these by ensuring rotational invariance, thus producing unique and repeatable signatures, as demonstrated in skin lesion classification studies [12].

In our previous work, we successfully implemented the Radial Fourier–Mellin and Hilbert transform to classify skin lesions accurately. This approach created stable image signatures across eight lesion types, producing vital classification metrics [12]. These results underscore the potential

of combining transform-based feature extraction with ML for robust lesion classification, especially when global invariance properties are required.

This study extends the previous approach by introducing the Hermite transform, which is applied independently to the image's red, green, and blue channels and grayscale. The Hermite transform is well-suited for capturing localized spatial features and intricate textures, essential for skin lesion analysis, as different color channels often contain unique structural information relevant to pigmentation, boundary definition, and textural detail [13–15]. Although the Hermite transform has primarily been used in areas such as biometric recognition, where fine-grained spatial analysis is essential, recent research suggests that the Hermite transform can be effectively applied to medical imaging for detailed feature extraction [16].

This study aims to enhance the classification of skin lesions by integrating the Hermite transform with the Radial Fourier–Mellin and Hilbert transform across separated RGB channels. We hypothesize that this hybrid approach will increase the robustness and accuracy of lesion classification, thereby advancing AI-driven, non-invasive diagnostic tools in dermatology. The following sections outline our methodology, experimental setup, results, and the impact of the Hermite transform on the overall classification performance.

The development of automated skin lesion classification systems has been an area of extensive research in medical image analysis, driven by the critical need for early skin cancer detection. Artificial intelligence (AI) and machine learning (ML) have significantly contributed to dermatology by improving the classification of lesion types using digital images. These approaches enhance diagnostic precision by employing feature extraction methods, multi-channel image analysis, and a range of mathematical transformations.

Several studies have highlighted the effectiveness of Fourier, Mellin, and Hilbert transforms for extracting stable invariant features under changes in rotation, scale, and translation—crucial properties for skin lesion analysis where lesions vary in appearance. The Fourier transform has been extensively utilized in medical imaging to extract global features, capturing frequency-based patterns in lesions for reliable classification [17]. The Mellin transform is often combined with the Fourier transform to improve accuracy by addressing size variations in lesions, introducing scale invariance, and enhancing the overall classification performance [18]. The Hilbert transform complements these by introducing rotational invariance, thus producing unique and repeatable signatures suitable for dermatological classification tasks, as shown in various studies on skin lesions [19].

Recent research has increasingly focused on hybrid approaches that combine multiple transforms to exploit complementary feature extraction strengths. For instance, in dermatology, hybrid techniques integrating texture descriptors, Fourier-based signatures, and ML classifiers have achieved superior classification performance over single-transform approaches by incorporating global and local features [12]. This integration has shown promise for skin lesions, where patterns such as edges, textures, and color gradients contribute significantly to lesion differentiation.
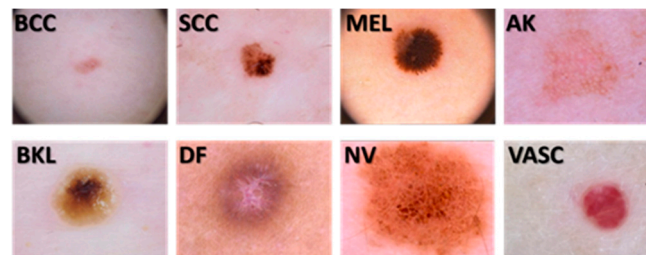
Though less commonly applied in dermatology, the Hermite transform has demonstrated potential in other fields, like facial recognition and fingerprint analysis, for its ability to capture fine-grained spatial and textural details. This transform uses Hermite polynomials to decompose images, effectively characterizing localized structures within complex image regions, such as lesion boundaries [15]. While few studies have applied the Hermite transform to multi-channel RGB images, evidence suggests that color channel-specific information enhances image analysis, especially for skin lesions where pigmentation and textural details are critical [20–22].

This study builds upon these advancements by integrating the Hermite transform with the Radial Fourier–Mellin and Hilbert transform across separated RGB channels to generate comprehensive feature sets for skin lesion classification. We aim to achieve higher accuracy and robust differentiation between lesion types by combining global invariant features and localized structures.

2. Materials and Methods

2.1. Image Dataset

Our dataset comprises digital images of various skin lesion types from the International Skin Imaging Collaboration (ISIC) 2019 dataset. Initial preprocessing involved the removal of images with significant noise (e.g., hair, artifacts) to enhance lesion visibility and ensure quality input for feature extraction. Data augmentation was applied to create a robust and balanced dataset, including rotations (at 45° increments) and scaling (100%, 95%, 90%, 85%, 80%). This preprocessing yielded a dataset of skin lesion images with each lesion type represented equally, reducing class imbalance and improving generalization for classification. Figure 1 illustrates some examples of skin lesion images contained in the dataset.

Figure 1. Some digital skin lesion images from dataset used in this work.

2.2. RGB Channel and Grayscale Separation

Each preprocessed image was separated into its red, green, and blue color channels and grayscale. This step allowed us to treat each channel independently, capitalizing on each channel's unique textural and color information. For each color channel, we computed transform-based features that emphasize distinct lesion characteristics across the RGB and grayscale spectrum. Separating the channels enabled us to retain and enhance color-specific information, which is vital for accurately distinguishing skin lesion types.

2.3. Hermite Transform

The Hermite transform is a specific type of polynomial transform and can be regarded as a model for image representation. This transformation serves as a method for signal decomposition and involves two main steps. Initially, the input signal $L(x, y)$ is combined with a window function through multiplication,

$$v(x - p, y - q), \tag{1}$$

at the positions p and q.

The goal is to achieve a comprehensive representation of the signal. This process is repeated at multiple positions spaced evenly across the image, forming a sampling grid S. At each (x, y), the pixel coordinates and the input signal are multiplied by the window function, while the original signal is

$$L(x, y) = \frac{1}{W(x, y)} \sum_{p,q \in S} L(x, y) v(x - p, y - q), \tag{2}$$

where,

$$W(x, y) = \sum_{p,q \in S} v(x - p, y - q) \tag{3}$$

is a weighting function.

The only requirement is that eq. (3) must be nonzero for all (x, y). Next, the signal within the window is expressed as a weighted sum of $G_{m,n-m}(x, y)$, with degrees m and $n - m$ relative to x, y, respectively. These polynomials are determined by the window function.

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} v^2(x, y) G_{m,n-m}(x, y) \times G_{l,k-l}(x, y) dx dy = \delta_{nl} \delta_{mk}. \tag{4}$$

Here, $n, l = 0, 1, 2, \dots, \infty$ and $m, k = 0, 1, 2, \dots, \infty$, while δ_{nl}, δ_{mk} represent the Kronecker delta functions, and $\times$ denotes element-wise multiplication. The process of converting the input signal into a weighted sum of polynomials, referred to as polynomial coefficients, is called the direct polynomial transform. These polynomial coefficients, $L_{m,n-m}(p, q)$, are obtained by convolving the original image with the analysis filters.

$$D_{m,n-m}(x, y) = G_{m,n-m}(x, y) v^2(-x, -y) \tag{5}$$

that is, for everything, $(p, q) \in S$,

$$L_{m,n-m}(p, q) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} L(x, y) \times D_{m,n-m}(x - p, y - q) dx dy \tag{6}$$

where,

$$G_{m,n-m}(p, q) = \frac{1}{\sqrt{2^n (n - m)! \, m!}} \times H_m \left(\frac{x}{\sigma} \right) H_{n-m} \left(\frac{y}{\sigma} \right), \tag{7}$$

$$v(x, y) = \frac{1}{\sqrt{\sqrt{\pi} \sigma}} e^{-\frac{x^2 + y^2}{2\sigma^2}}, \tag{8}$$

where $v(x, y)$ is Gaussian window function, σ is the standard deviation of the Gaussian window function and,

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}, \quad n = 0, 1, 2, \dots \tag{9}$$

where $H_n(x)$ is the n-th Hermite polynomial.

Using the convolution form in eq. (6) it is defined as,
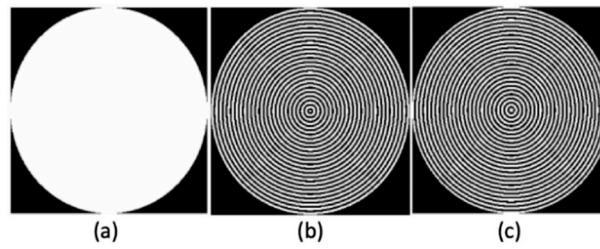
$$L_{m,n-m}(p, q) = L(x, y) * D_{m,n-m}(x, y). \tag{10}$$
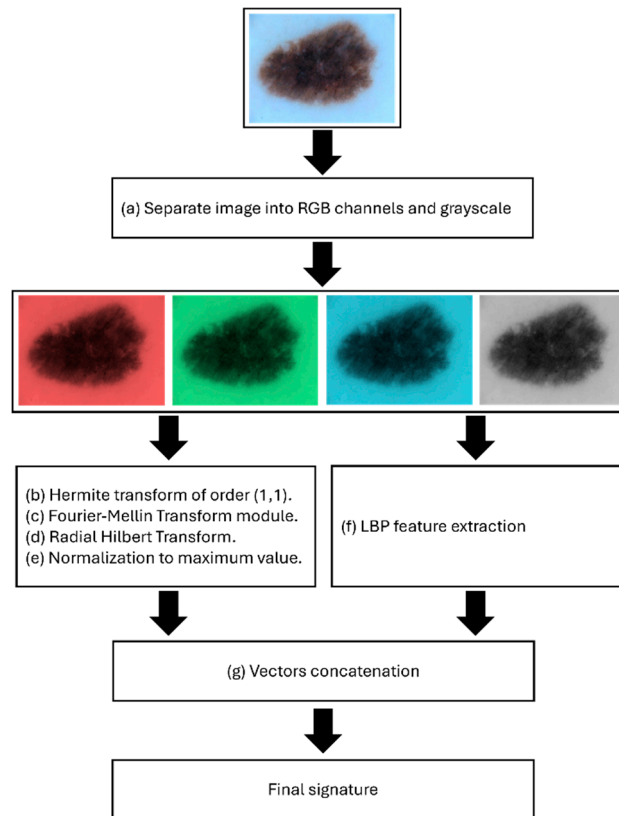
2.4. The Signatures

We generated multiple signature vectors for the Hermite transform of order (1,1) for each RGB channel and grayscale images using a dataset of 362,680 samples created through data augmentation. The augmentation process included five scaling percentages (100%, 95%, 90%, 85%, and 80%) and eight rotation angles (45°, 90°, 135°, 180°, 225°, 270°, 315°, and 360°). These descriptors utilized invariance properties to translation and scaling from the Fourier and Mellin transforms, respectively. The Hilbert transform was applied for rotational invariance. Unique image signatures were computed by summing pixel values within each ring produced by Hilbert masks used as filters. The texture descriptors were then incorporated into the previously generated radial Fourier signatures.

This process resulted in a one-dimensional representation or signature of the skin lesion digital image, as illustrated in Figures 2 and 3. The original image $Im(x, y)$ consists of three RGB matrix channels (red, green, and blue). These were separated into their respective primary color channels for the application of the radial Fourier–Mellin method and uniform Local Binary Pattern (LBP) feature extraction. Additionally, the grayscale skin lesion image was derived as a weighted sum of RGB values using the formula 0.299R + 0.587G + 0.114B.

Figure 2. (a) Binary disk. (b) H_R mask. (c) H_I mask.

Figure 3. Methodology proposed for image embeddings generation.

2.5. Radial Fourier-Mellin Signatures Through Hilbert Transform

To create the radial Fourier–Mellin signatures, the image was first separated into its RGB channels and grayscale (Figure 3(a)). After that, the Hermite transform of order (1,1) is obtained from each image component, RGB channels and grayscale (Figure 3(b)).

Next, the magnitude of the Fourier–Mellin (FM) transform for each skin lesion digital image, denoted as $Im(x, y)$, was computed using the following equation (Figure 3(c)).

$$|F_M(s,t)| =$$
$$\iint_0^\infty |FT[Im(x,y)]| x^{(s-1)} y^{(t-1)} \mathrm{dxdy} \tag{11}$$
$$= M\{|FT[Im(x,y)]|\},$$

here, $|F_M(s,t)|$ represents the magnitude of the Mellin transform, which provides scale invariance for objects in the image. This is essential because the skin lesion digital images were captured at varying distances between the lesion and the camera. As a result, lesions appear smaller at greater distances and larger when the distance is shorter. The coordinates (s,t) represent the 2D positions of the transformed (x,y) pixel coordinates on Mellin's plane. These original (x,y) coordinates correspond to the magnitude of the Fourier transform of the image, $|FT[Im(x,y)]|$, leveraging its

translation invariance. Consequently, at this stage, the object (skin lesion) in the image becomes invariant to both translation and scale.

By applying the Hilbert transform, rotational invariance of the skin lesion in the image is also achieved. The Hilbert transform of the image is expressed as:

$$\mathcal{F}\{H_r[Im(x,y)]\} =$$
$$e^{ip\theta}FT[Im(x,y)] = \qquad (12)$$
$$e^{ip\theta}F(u,v),$$

here, p represents the order of the radial Hilbert transform, and θ is the angle in the frequency domain corresponding to the pixel coordinates (x,y) after their transformation into Fourier plane coordinates (u,v). This angle is calculated as $\theta = \text{aco s}\left(u/\sqrt{u^2+v^2}\right)$.

Using Euler's formula, binary ring masks were generated for the RGB channels and the grayscale skin lesion digital image. These masks utilized both the real H_R and imaginary H_I components of the radial Hilbert transform of the image (Figure 2).

$$H_R = Re[H_r(u,v)] = \begin{cases} 1, & if \ sin(p\theta) > 0 \\ 0, & otherwise \end{cases}, \qquad (13)$$
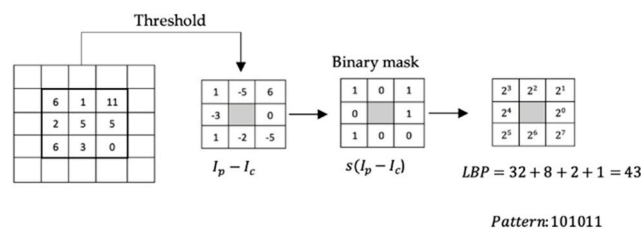
$$H_I = Im[H_r(u,v)] = \begin{cases} 1, & if \ cos(p\theta) > 0 \\ 0, & otherwise \end{cases}. \qquad (14)$$

The binary ring masks generated earlier were used to filter the skin lesion digital images that had been processed with the magnitude of the Fourier–Mellin transform (Figure 3(c)). The process involved summing the pixel values within each ring, resulting in two distinct signatures for each grayscale skin lesion image $Sgray_{H_R}$ and $Sgray_{H_I}$; and for its RGB channels: SR_{H_R}, SR_{H_I}, SG_{H_R}, SG_{H_I}, SB_{H_R}, and SB_{H_I} (Figure 3 (d)). Finally, each signature is normalized by its maximum value (Figure 3 (e)).

To incorporate texture descriptors, we applied the uniform Local Binary Pattern (LBP) technique (Figure 3(f)), a widely used tool in computer vision and image processing for texture analysis.

LBP is a simple yet effective descriptor that captures textures, edges, corners, spots, and flat regions. For each 3×3 pixel block, the intensity of the eight surrounding pixels is compared to the intensity of the central pixel, which serves as the threshold. If the intensity of a neighboring pixel is greater than or equal to the central pixel, its position is assigned a value of 1; otherwise, it is assigned 0. After comparing all pixels, a binary sequence is formed.

This binary sequence is then converted into a decimal value by multiplying each position by its corresponding weight (decimal value) and summing the results. The final LBP value is used to label the central pixel. Figure 4 illustrates the LBP calculation for a pixel with $P = 8$ neighboring pixels.

Figure 4. Procedure to calculate LBP.

To compute the Local Binary Pattern (LBP) for a grayscale image, the following equation is used:

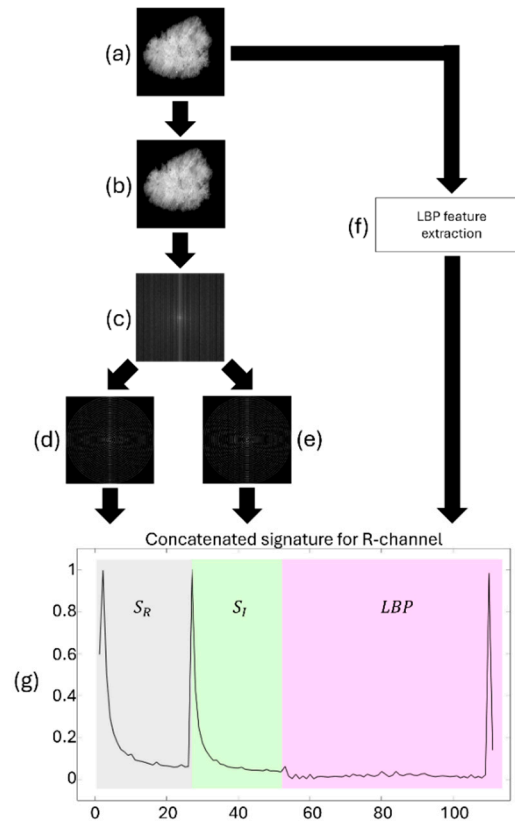$$LBP(x_c, y_c) = \sum_{p=0}^{P-1} s(I_p - I_c)2^p, \qquad (15)$$

where (x_c, y_c) represents the central pixel's coordinates, P is the number of neighboring pixels, I_P is the intensity of the neighboring pixel, I_c is the intensity of the central pixel, and s is a step function defined as:

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & another\ value \end{cases} . \tag{16}$$

The uniform LBP (LBP-U) is a variant of the standard LBP that reduces the dimensionality of the characteristic vector and provides rotational invariance. An LBP pattern is considered uniform if it has at most two transitions between 0 and 1 in the binary sequence. For example, patterns like 11111111 (0 transitions), 11111000 (1 transition), and 11001111 (2 transitions) are uniform, whereas 11010110 (6 transitions) and 11001001 (4 transitions) are non-uniform.

In an eight-pixel neighborhood, 256 patterns can be generated, of which 58 are uniform. These uniform patterns are assigned unique labels (1–58), while all non-uniform patterns are grouped under a single label (59). This work utilizes the LBP-U technique.

After calculating the uniform LBP for each pixel, a histogram of LBP values is constructed to represent the texture distribution for both RGB and grayscale images LBP_R, LBP_G, LBP_B, and LBP_{Gray}. These histograms are then concatenated to form 444 components, creating one-dimensional object signatures (Figure 3(g)). Figure 5 exemplify this procedure applied to R chanel.

Figure 5. (a) Processed R-channel image, (b) Hermite transform of order (1,1), (c) Fourier-Mellin transform modulus, (d) HR part of the radial Hilbert transform, (e) HI part of the radial Hilbert transform, (f) LBP feature extraction, (g) Concatenated signature for R-channel.

2.6. Signatures Classification

We implemented a neural network model for classification using the Keras Sequential in Python. The model was structured to process the 444-dimensional feature vectors generated by the radial Fourier-Mellin signatures and LBP histogram from each RGB channel and grayscale. This neural network was designed with multiple dense layers to capture complex relationships within the feature space.

2.7. Model Architecture

The neural network architecture consists of six dense layers. The first layer has 100 units with a ReLU activation function and takes the 444-dimensional feature vector as input. Each four hidden layer has 100 units with ReLU activation to introduce non-linearities and model complex patterns. The final layer has eight units with a softmax activation function, corresponding to the classification of the eight skin lesion types in the dataset.

2.8. Model Compilation

The model was compiled using the Adam optimizer, which is well-suited for this multi-layer architecture due to its adaptive learning rate capabilities. The loss function was set to sparse categorical cross-entropy, suitable for multi-class classification with integer-labeled target classes. Model accuracy was tracked as a performance metric during training.

2.9. Training Procedure

The model was trained on the feature set using 400 epochs, with x_{train} as the feature input and y_{train} as the target lesion labels. This training process allowed the model to learn representations of the different lesion types based on the extracted features.

2.10. Model Performance

The neural network model's performance was evaluated using accuracy as the primary metric on both the training and test sets. Additionally, to ensure robustness, we calculated and reported recall, FP rate, specificity, precision, accuracy, and F1 score for each lesion class, providing a comprehensive assessment of the model's classification capabilities across lesion types.

3. Results

We randomly selected images to classify. The classes in the dataset were balanced, with 1840 images used for each type of skin lesion to prevent classification bias. Then, the deep learning model described in the "Classification" section was trained using a data split of 30% to test and 70% to train.

To assess the classification performance of the proposed methodology, a variety of standard metrics were employed. Recall (Eq. 17) quantifies the proportion of actual positives correctly identified by the model, emphasizing its ability to capture all instances of a given class. False Positive Rate (Eq. 18) reflects the proportion of negative instances misclassified as positive, serving as a complement to specificity. Specificity (Eq. 19) measures the model's capability to correctly identify negative instances, calculated as one minus the FP rate. Precision (Eq. 20) indicates the percentage of correctly predicted positive instances, demonstrating the reliability of positive predictions. Accuracy (Eq. 21) provides an overall evaluation of the model, representing the ratio of correctly classified instances (both positive and negative) to the total number of instances. Lastly, the F1 score (Eq. 22) offers a balanced measure of precision and recall, expressed as their harmonic mean, making it particularly useful in scenarios with imbalanced datasets. Collectively, these metrics provide a comprehensive view of the model's classification performance.

$$Recall(C_i) = \frac{TP_i}{TP_i + FN_i}, \tag{17}$$

$$FP\ rate(C_i) = \frac{FP_i}{FP_i + TN_i}, \tag{18}$$

$$Specificity(C_i) = 1 - FP\ rate(C_i), \tag{19}$$

$$Precision(C_i) = \frac{TP_i}{TP_i + FP_i}, \tag{20}$$

$$accuracy(C_i) = \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i},$$ (21)

$$F1\ score(C_i) = \frac{2 \cdot Precision(C_i) \cdot Recall(C_i)}{Precision(C_i) + Recall(C_i)},$$ (22)

where:

TP_i: true positives for class i.

TN_i: true negatives for class i.

FP_i: false positives for class i.

FN_i: false negatives for class i.

The confusion matrix, Figure 6, reveals that the classification model performs well overall, with diagonal solid dominance indicating accurate predictions for most classes. High accuracy is observed for vascular lesions (VASC), dermatofibroma (DF), squamous cell carcinoma (SCC), and actinic keratosis (AK), with minimal misclassifications. However, notable confusion exists between melanocytic nevus (NV) and melanoma (MEL), as well as between MEL and basal cell carcinoma (BCC), suggesting overlapping features among these lesion types. Additionally, benign keratosis (BKL) is occasionally misclassified as NV.

Figure 6. Confusion matrix obtained on test set.

The ROC curves in Figure 7 illustrate the performance of the multiclass classification model across eight skin lesion classes, showing the relationship between the actual positive rate (sensitivity) and the false positive rate. The area under the curve (AUC) values indicates high classification performance for all classes. Actinic keratosis (AK) and dermatofibroma (DF) achieved perfect classification with an AUC of 1.00, indicating the model's ability to distinguish these classes with no errors. Similarly, vascular lesions (VASC) and squamous cell carcinoma (SCC) demonstrated near-perfect separability with AUC values of 0.99. Benign keratosis (BKL) and basal cell carcinoma (BCC) followed by AUC values of 0.94 and 0.95, respectively, reflecting strong performance despite minor overlaps. Melanoma (MEL) and melanocytic nevus (NV) exhibited slightly lower AUC values of 0.92 and 0.91, suggesting some challenges in differentiation. The micro-average AUC of 0.97 confirms robust overall performance across all classes, underscoring the model's ability to handle the complexities of multiclass skin lesion classification effectively.
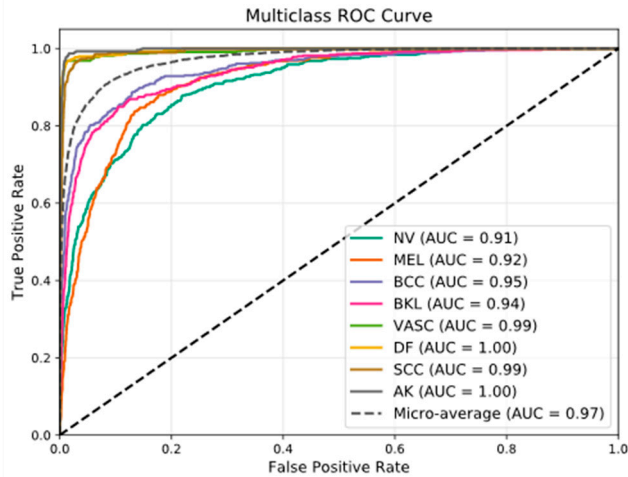
Figure 7. Multiclass ROC curve.

Table 1. Performance metrics multiplied by 100.

Class	Recall	FP rate	Specificity	Precision	Accuracy	F1 Score
NV	55.85 ± 2.33	4.50 ± 0.33	95.50 ± 0.33	63.92 ± 1.43	90.60 ± 0.22	59.34 ± 1.44
MEL	59.31 ± 2.11	5.25 ± 0.44	94.75 ± 0.44	62.07 ± 1.57	90.34 ± 0.32	60.39 ± 1.26
BCC	75.16 ± 1.18	2.93 ± 0.29	97.07 ± 0.29	78.54 ± 1.76	94.35 ± 0.29	76.73 ± 1.16
BKL	71.68 ± 1.71	3.94 ± 0.31	96.06 ± 0.31	72.61 ± 1.32	92.99 ± 0.28	72.01 ± 1.06
VASC	94.77 ± 0.99	0.97 ± 0.10	99.03 ± 0.10	93.37 ± 0.60	98.50 ± 0.12	94.03 ± 0.50
DF	97.54 ± 0.72	1.17 ± 0.13	98.83 ± 0.13	92.29 ± 0.79	98.66 ± 0.17	94.83 ± 0.65
SCC	97.01 ± 0.83	1.59 ± 0.17	98.41 ± 0.17	89.71 ± 1.00	98.23 ± 0.22	93.20 ± 0.82
AK	98.62 ± 0.58	0.98 ± 0.16	99.02 ± 0.16	93.64 ± 0.95	98.97 ± 0.18	96.05 ± 0.65
Overall Average	81.24 ± 2.19	2.67 ± 0.22	97.33 ± 0.22	80.77 ± 1.64	95.33 ± 0.45	80.82 ± 1.90

The entire workflow was repeated 30 times to account for the inherent randomness in processes such as balancing class samples, splitting the dataset into training and testing sets, and the stochastic nature of neural network training. This repetition ensures that the results are not biased by any random data partitioning or model initialization instance. By performing the process 30 times, the average of the performance metrics (e.g., accuracy, recall, precision) could be calculated, along with their error intervals at a 95% confidence level. This approach provides robust and statistically reliable metrics, minimizing the impact of random fluctuations and offering a more accurate assessment of the model's performance.

The resulting metrics, Table 1, demonstrate the model's performance. Recall values ranged from 55.85% ± 2.33 for melanocytic nevus (NV) to 98.62% ± 0.58 for actinic keratosis (AK), showing varying sensitivity in correctly identifying different lesion types. Similarly, false positive (FP) rates were lowest for AK (0.98% ± 0.16) and vascular lesions (VASC) (0.97% ± 0.10), reflecting high precision in these classes.

Specificity remained consistently high across all classes, ranging from 94.75% ± 0.44 for melanoma (MEL) to 99.03% ± 0.10 for VASC, indicating the model's strong ability to identify true negatives. Precision followed a similar trend, with higher values for classes like AK (93.64% ± 0.95) and VASC (93.37% ± 0.60) and slightly lower for NV (63.92% ± 1.43) and MEL (62.07% ± 1.57).

Overall accuracy averaged 95.33% ± 0.45, with AK, SCC, and DF achieving the highest individual accuracies (above 98%). The F1 score, reflecting the harmonic mean of precision and recall, varied from 59.34% ± 1.44 for NV to 96.05% ± 0.65 for AK. The overall F1 score was 80.82% ± 1.90, highlighting the model's balanced performance across all classes.

4. Discussion

The results of this study confirm the effectiveness of the proposed hybrid approach for skin lesion classification. By integrating the Hermite transform with the Radial Fourier–Mellin and Hilbert transform, we achieved a balanced extraction of local and global features essential for robust classification. The Hermite transform's capability to decompose images into localized polynomial coefficients allowed us to analyze fine-grained spatial details within skin lesions, particularly in the RGB channels. This analysis is critical for identifying pigmentation, texture variations, and boundary definitions, which are vital in distinguishing benign lesions from malignant ones.

On the other hand, the Radial Fourier–Mellin transform ensured scale and translation invariance, addressing challenges posed by varying lesion sizes and positions across images. The Hilbert transform further complemented these by introducing rotational invariance, thus generating repeatable and stable signatures. Combining these three transforms provided a rich and comprehensive feature set, enhancing the discriminative power of the machine learning models applied.

The multi-channel analysis (RGB and grayscale) proved particularly beneficial. Each channel emphasized unique lesion characteristics, with the red channel highlighting pigmentation, the green channel focusing on intermediate structures, and the blue channel enhancing fine textures. This approach aligns with recent findings where multi-channel feature extraction outperformed single-channel medical image analysis methods.

Table 2 presents a comparative analysis of various state-of-the-art models for skin lesion classification using different datasets. Among the listed models, Inception-v2 achieved the highest recall (0.9015) and F1 score (0.8876), indicating strong sensitivity and overall performance. VGG19 and ResNeXt101 also demonstrated competitive results across precision, accuracy, and F1 score. However, it is important to highlight that our proposed method shows strength in terms of false positive (FP) rate and specificity, with values of 0.0267 and 0.9733, respectively—metrics not reported for the other methods. This low FP rate suggests a significantly reduced rate of incorrect positive classifications, which is crucial in medical applications to avoid unnecessary patient anxiety or treatment. Despite a slightly lower recall (0.8124), the method achieves the highest accuracy (0.9533) among all models compared, indicating excellent overall classification reliability. These results demonstrate that our approach provides a more balanced trade-off between minimizing false alarms and maintaining high classification performance.

Table 2. Comparison with state of art.

Model	Dataset	Recall	FP rate	Specificity	Precision	Accuracy	F1 Score
2D superpixels + RCNN [23]	HAM-10000	0.8450	-	-	0.8349	0.8550	0.8530
ResNeXt101 [24]	ISIC-2019	0.8810	-	-	0.8740	0.8850	0.8830
MobileNetV2 [25]	ISIC-2019	0.8633	-	-	0.7890	0.8530	-
VGG19 [26]	ISIC-2019, Derm-IS	0.8666	-	-	0.9070	0.8857	0.8765
ConvNet [27]	ISIC-2019, Derm-IS	0.8747	-	-	0.8614	0.8690	-
Inception-v2 [28]	ISIC-2019	0.9015	-	-	0.8737	0.8904	0.8876
This work	ISIC-2019	0.8124	0.0267	0.9733	0.8077	0.9533	0.8082

While the results are promising, some limitations exist. The computational complexity of the Hermite transform, mainly when applied to multiple channels and high-resolution images, can be a challenge for real-time deployment. Future studies could explore optimizing this process or applying dimensionality reduction techniques to streamline computation. Testing the method on other datasets or clinical settings will also validate its generalizability and real-world applicability.

This study presents a novel hybrid approach for automated skin lesion classification, combining the Hermite transform, Radial Fourier–Mellin transform, and Hilbert transform. Integrating these transforms across RGB and grayscale channels allows for comprehensive feature extraction,

capturing both global invariance properties and localized spatial details. Experimental results demonstrate that this method enhances classification, accuracy and robustness, outperforming traditional transform-based approaches.

The findings underscore the potential of hybrid feature extraction techniques for advancing AI-driven dermatological diagnosis tools. Future work will focus on reducing computational costs and expanding the model's validation to other datasets and clinical applications. By improving the precision of automated skin lesion classification, this method contributes to non-invasive, early detection strategies, ultimately benefiting patient care and clinical decision-making.

References

1. Shetty, B.; Fernandes, R.; Rodrigues, A.P.; Chengoden, R.; Bhattacharya, S.; Lakshmanna, K. Skin lesion classification of dermoscopic images using machine learning and convolutional neural network. Sci. Rep. 2022, 12, 18134.

2. Wu, Y.; Chen, B.; Zeng, A.; Pan, D.; Wang, R.; Zhao, S. Skin cancer classification with deep learning: a systematic review. Front. Oncol. 2022, 12, 893972.

3. Dubal, P.; Bhatt, S.; Joglekar, C.; Patil, S. Skin cancer detection and classification. In Proc. 6th Int. Conf. Electr. Eng. Informatics (ICEEI), Nov. 2017; pp. 1–6.

4. Benjamin, J.R.; Jayasree, T. Improved medical image fusion based on cascaded PCA and shift invariant wavelet transforms. Int. J. Comput. Assist. Radiol. Surg. 2018, 13, 229–240.

5. Barajas-Garcia, C.; Solorza-Calderon, S.; Gutierrez-Lopez, E. Scale, translation and rotation invariant wavelet local feature descriptor. Appl. Math. Comput. 2019, 363, 124594.

6. Kim, H.K.; Kim, J.D. Region-based shape descriptor invariant to rotation, scale and translation. Signal Process. Image Commun. 2000, 16, 87–93.

7. Popa, C.A.; Cernăzanu-Glăvan, C. Fourier transform-based image classification using complex-valued convolutional neural networks. In Advances in Neural Networks–ISNN 2018: 15th International Symposium on Neural Networks, Minsk, Belarus, Jun. 25–28, 2018; Springer International Publishing: Cham, Switzerland, 2018; Vol. 15, pp. 300–309.

8. Uzun, I.S.; Amira, A.; Bouridane, A. FPGA implementations of fast Fourier transforms for real-time signal and image processing. IEE Proc. Vis. Image Signal Process. 2005, 152, 283–296.

9. Braccini, C.; Gambardella, G.; Grattarola, A. Digital image processing by means of generalized scale-invariant filters. In Issues in Acoustic Signal–Image Processing and Recognition; Springer Berlin Heidelberg: Berlin, Germany, 1983; pp. 315–329.

10. De Sena, A.; Rocchesso, D. A fast Mellin and scale transform. EURASIP J. Adv. Signal Process. 2007, 2007, 1–9.

11. Bigot, J.; Gamboa, F.; Vimond, M. Estimation of translation, rotation, and scaling between noisy images using the Fourier–Mellin transform. SIAM J. Imaging Sci. 2009, 2, 614–645.

12. Guerra-Rosas, E.; López-Ávila, L.F.; Garza-Flores, E.; Vidales-Basurto, C.A.; Álvarez-Borrego, J. Classification of skin lesion images using artificial intelligence methodologies through radial Fourier–Mellin and Hilbert transform signatures. Appl. Sci. 2023, 13, 11425.

13. Garza-Flores, E.; Guerra-Rosas, E.; Álvarez-Borrego, J. Spectral indexes obtained by implementation of the fractional Fourier and Hermite transform for the diagnosis of malignant melanoma. Biomed. Opt. Express 2019, 10, 6043–6056.

14. Martins, A.S.; et al. A Hermite polynomial algorithm for detection of lesions in lymphoma images. Pattern Anal. Appl. 2021, 24, 523–535.

15. Barba-J, L.; Vargas-Quintero, L.; Calderón-Agudelo, J.A. Bone SPECT/CT image fusion based on the discrete Hermite transform and sparse representation. Biomed. Signal Process. Control 2022, 71, 103096.

16. Agrafioti, F.; Gao, J.; Hatzinakos, D.; Yang, J. Heart biometrics: Theory, methods and applications. Biometrics 2011, 3, 199–216.

17. Damian, F.A.; Moldovanu, S.; Dey, N.; Ashour, A.S.; Moraru, L. Feature selection of non-dermoscopic skin lesion images for nevus and melanoma classification. Computation 2020, 8, 41.

18. McGuire, M. An image registration technique for recovering rotation, scale and translation parameters. NEC Res. Inst. Tech. Rep. 1998, TR-98-018.

19. Ngo, L.H.; Luong, M.; Sirakov, N.M.; Viennet, E.; Le-Tien, T.T. Skin lesion image classification using sparse representation in quaternion wavelet domain. Signal Image Video Process. 2022, 16, 1721–1729.

20. Was, L.; Milczarski, P.; Stawska, Z.; Wyczechowski, M.; Kot, M.; Wiak, S.; et al. Analysis of dermatoses using segmentation and color hue in reference to skin lesions. In Artificial Intelligence and Soft Computing: 16th International Conference, ICAISC 2017, Zakopane, Poland, Jun. 11–15, 2017; Springer International Publishing: Cham, Switzerland, 2017; Vol. 16, pp. 677–689.

21. Gerstenblith, M.R.; Shi, J.; Landi, M.T. Genome-wide association studies of pigmentation and skin cancer: A review and meta-analysis. Pigment Cell Melanoma Res. 2010, 23, 587–606.

22. Gallagher, R.P.; Hill, G.B.; Bajdik, C.D.; Coldman, A.J.; Fincham, S.; McLean, D.I.; Threlfall, W.J. Sunlight exposure, pigmentation factors, and risk of nonmelanocytic skin cancer: II. Squamous cell carcinoma. Arch. Dermatol. 1995, 131, 164–169.

23. Peng, S.; Jiang, H.; Wang, H.; Alwageed, H.; Zhou, Y.; Sebdani, M.M.; Yao, Y.D. Modulation classification based on signal constellation diagrams and deep learning. IEEE Trans. Neural Netw. Learn. Syst. 2019, 30, 718–727.

24. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal loss for dense object detection. IEEE Trans. Pattern Anal. Mach. Intell. 2020, 42, 318–327.

25. Malik, H.; Farooq, M.S.; Khelifi, A.; Abid, A.; Qureshi, J.N.; Hussain, M. A comparison of transfer learning performance versus health experts in disease diagnosis from medical imaging. IEEE Access 2020, 8, 139367–139386.

26. Mijwil, M.M. Skin cancer disease images classification using deep learning solutions. Multimed. Tools Appl. 2021, 80, 26255–26271.

27. Indraswari, R.; Rokhana, R.; Herulambang, W. Melanoma image classification based on MobileNetV2 network. Procedia Comput. Sci. 2022, 197, 198–207.

28. Zhi, S.; Li, Z.; Yang, X.; Sun, K.; Wang, J. A multiclassification model for skin diseases using dermatoscopy images with Inception-v2. Appl. Sci. 2024, 14, 10197.