مجلة جامعة بني وليد للعلوم الإنسانية والتطبيقية
Bani Waleed University Journal of Humanities and Applied Sciences
تصدر عن– جامعة بني وليد – ليبيا
Website:https://jhas-bwu.com/index.php/bwjhas/index

المجلد العاشر ــ العدد الرابع ــ 2025 الصفحات (20- 33) ISSN3005-3900

A Deep Learning-Based Method for Skin Cancer Detection

Rabha O. AbdElsalam [1]* , Safa Abdelkarem Elgzali [2], Sahar Q. Saleh [3]

[1.2] Department of Computer Science, College of Humanities and Applied Sciences, University of Benghazi, Tokra, Libya.

[3] Department of Computer Science, Taiz University, Taiz, Yemen.

rabha.omar@uob.edu.ly

طريقة تعتمد على التعلم العميق للكشف عن سرطان الجلد

رابحة عمر عبدالسلام [1] *، صفاء عبدالكريم الغزالي [2]، سحر قاسم صالح [3]

[1.2] قسم تقنية المعلومات، كلية العلوم الانسانية والتطبيقية، جامعة بنغازي،توكرة، ليبيا.

[3] قسم علوم حاسوب، كلية العلوم، جامعة تعز، تعز، اليمن.

تاريخ الاستلام: 2025-07-10 تاريخ القبول: 2025-08-02 تاريخ النشر : 2025-10-01

الملخص:

يُعد سرطان الجلد من الأمراض الشائعة التي قد تهدد الحياة، الأمر الذي يجعل الحاجة ملحة لوجود طرق دقيقة للكشف المبكر عنه. في هذه الدراسة، تم توظيف تقنيات التعلم العميق لمعالجة التحدي المتمثل في تصنيف الآفات الجلدية إلى فئتين: حميدة أو خبيثة. ويقترح البحث إطاراً متكاملاً يجمع بين تحويل الموجات المتقطع (DWT) لاستخراج الخصائص المكانية والترددية، وتحليل المكونات الرئيسية (PCA) لتقليل الأبعاد، إضافة إلى الشبكات العصبية العميقة المتكررة (RNN) المدعومة بطبقات LSTM لأغراض التصنيف. تتم معالجة صور الآفات الجلدية عبر تحليلها بـ DWT لاستخراج السمات المنخفضة التردد الأكثر تميزاً، ثم تطبيق PCA للتخلص من التكرار وتحسين كفاءة المعالجة. بعد ذلك تُمرَّر هذه السمات إلى نموذج RNNقادر على تمثيل العلاقات المعقدة داخل البيانات.

أظهرت النتائج قدرة قوية على التعميم، حيث حقق النموذج دقة متوسطة بلغت 96.33% وقيمة F1-score = 0.9803 عبر جميع الطيات. إن التكامل بين DWT و PCA و RNN ساهم في بناء نظام تشخيصي عالي الكفاءة والدقة، مما يبرز إمكاناته الواعدة في التطبيقات العملية بمجال تحليل الصور الطبية والكشف المبكر عن سرطان الجلد.

الكلمات الدالة: تحويل الموجات المنفصلة، تحليل المكونات الأساسية، الشبكات العصبية المتكررة.

Abstract

Skin cancer is a common and potentially fatal disease, necessitating accurate and early detection methods. This study leverages the power of deep learning to address the challenge of classifying skin lesions into benign or malignant categories. An integrated framework is proposed, combining Discrete Wavelet Transform (DWT) for spatial and frequency-based

feature extraction, Principal Component Analysis (PCA) for dimensionality reduction, and Recurrent Neural Networks (RNN) with Long Short-Term Memory (LSTM) layers for classification. The system processes lesion images by decomposing them using DWT to extract salient low-frequency features, followed by PCA to eliminate redundancy and enhance computational efficiency. These refined features are then analyzed through an RNN architecture capable of modeling complex dependencies within the data.
The results demonstrate strong generalization performance, with the model achieving an average accuracy of 96.33% and an F1-score of 0.9803 across folds. The synergy between DWT, PCA, and RNN contributes to a highly efficient and accurate diagnostic system, underscoring its potential for real-world applications in medical image analysis and early detection of skin cancer.

Keywords: Discrete Wavelet Transform; Principal Component Analysis; Recurrent Neural Networks.

1. Introduction

Skin cancer is one of the most common types of cancer worldwide, starting in the skin cells. It is a widespread malignant tumor and often appears on sun-exposed skin. The three main types of skin cancer include melanoma, squamous cell carcinoma, and basal cell carcinoma. The cancer is characterized by the ability of its cells to grow abnormally, leading to the formation of tumors that can spread quickly if not detected in their early stages. The most prominent signs of skin cancer are abnormal changes in the appearance of the skin, such as the appearance of new moles or changes in existing moles, which are often warning signs of the presence of skin cancer. Early detection of this type of cancer is crucial to improving survival rates and reducing the risk of progression to later stages. However, doctors face significant challenges in identifying skin tumors in their early stages due to the inconspicuous nature of symptoms at these stages. Therefore, Artificial intelligence (AI) has emerged as a powerful tool to accelerate and enhance diagnostic accuracy. AI systems can analyze medical images of skin lesions and detect subtle patterns and indicators often missed by the human eye.

These systems process vast amounts of medical data, uncovering complex relationships that traditional methods might overlook. Early detection of skin cancer can result in improved treatment outcomes and more effective therapies. In general, AI is driving major advances in skin cancer detection and treatment [1] [2].
Esteva et al. (2017) conducted groundbreaking research by using CNN-based model for skin cancer detection, which analyzed photos of skin lesions. The model, trained on a dataset of over 129,000 skin images, achieved an accuracy rate of 85% [3]. Ichim et al. (2020) developed a skin lesion detection system using deep learning techniques, particularly CNN models such as GoogleNet, ResNet-101, and NasNet-Large, combined with an SVM based classification method. The system achieved an overall classification accuracy of 93%, effectively distinguishing between benign tumors and melanoma [4]. Won et al. (2021) developed a computer-aided diagnostic system that integrates a CNN-based classification model with a U-Net segmentation model to differentiate malignant melanoma from benign skin tumors using RGB skin images. This approach achieved an accuracy of 80.06% for melanoma classification and an 81.1% Dice similarity coefficient for lesion segmentation [5]. Additionally, in 2021, Jain et al. utilized the HAM10000 dataset to classify multi-class skin cancer using six transfer learning networks. The Xception network outperformed the others, achieving a classification accuracy of 90.48% [6]. Hussain et al. (2022) introduced a CNN model based on Inception-v3, achieving an 88% success rate in melanoma detection. Their study emphasized the importance of preprocessing and noise reduction to enhance image quality, which significantly aided the model in extracting

discriminative features [7]. Hussain et al. SkinLesNet, a deep model based on CNN designed to classify seborrheic keratosis, nevi, and melanoma using high-resolution data from various datasets such as PAD-UFES-20, HAM10000, and ISIC2017 [8].

In 2023, Bhatti et al. achieved an 88% accuracy rate in skin cancer categorization using deep neural networks (DNNs). Their research demonstrated the efficiency of DNNs in processing complex medical imaging data by testing the models on large datasets of skin cancer [9]. Additionally, Ansari et al. (2023) developed a skin cancer detection model using DNNs based on ResNet and Inception architectures. This model achieved an accuracy of 94.6%, comparable to the diagnostic performance of professional dermatologists, highlighting its potential to enhance early detection of melanoma and other skin cancers [10]. In 2023 Jitendra V. Tambourine et al. conducted a study aiming to enhance skin cancer detection, particularly melanoma, by integrating machine learning and deep learning techniques. The proposed model utilized deep neural networks to extract automatic features from images, alongside handcrafted features derived using methods such as Contourlet Transform and Local Binary Pattern Histogram. The results showed that the model achieved a classification accuracy of 93%, outperforming traditional methods and expert dermatologists [11]. Another research 2024, Claret et al. further refined this approach by combining CNN and DWT achieving the same classification accuracy of 94%. This method surpassed conventional techniques, solidifying its potential for early detection of melanoma [12].

Recurrent Neural Network (RNN) is a deep learning technique, which has been widely used in medical applications such as cancer detection. Aziz et al. (2018) evaluated Temporal Enhanced Ultrasound (TeUS) data for prostate cancer diagnosis in 2018 using RNN, more especially Long Short-Term Memory (LSTM) networks. The study achieved a high accuracy of 93% and comprised data from 255 biopsy samples and 157 individuals. Comparing temporal modeling using RNN to conventional techniques, the study showed that the former greatly increases the accuracy of early prostate cancer diagnosis [13]. Ganguly et al. (2023) devised a Generative RNNs (GRNNs) framework that develops accurate and meaningful synthetic data of patients with metastatic cancer after surgery. The goal of this framework is to address the challenges of data sharing by generating synthetic datasets that preserve the statistical properties of the original data while ensuring patient anonymity. The model was used on167,474 patients and it achieved 93.2% accuracy at a loss of 0.21% [14]. Zareen et al. (2024). Achieved 94.48% accuracy rate in skin cancer classification using a hybrid deep learning model combining Convolutional Neural Networks (CNNs) and (RNNs). Their research demonstrated the efficiency of utilizing ResNet-50 architecture for spatial feature extraction and STM layer for temporal analysis, tested on the ISIC dataset to classify nine types of skin cancer [15].

Soglia (2025) developed an advanced AI model based on LC-OCT images to classify melanocytic skin lesions into three categories: benign nevi, atypical/dysplastic nevi, and melanoma. The study used retrospective analysis of vertical DICOM images enhanced with several filters to improve the signal-to-noise ratio. Pixel clustering techniques and biomarker extraction were applied using specialized software. Machine learning models including logistic regression, decision trees, and random forests were trained to differentiate melanoma from other nevi. While the models showed high accuracy on training data, they exhibited over fitting, resulting in a decreased accuracy of approximately 70–75% on independent test data. This study was the first to demonstrate the feasibility of in-vivo melanoma classification using LC-OCT combined with AI [26].

Hosseinzadeh et al. (2025) proposed a novel skin cancer diagnosis model combining deep learning and ensemble learning methods, published in PLOS ONE. The model integrated outputs from multiple deep learning models trained on skin lesion datasets, improving classification accuracy between benign and malignant lesions. The ensemble model outperformed individual

models, achieving an accuracy close to 94%, while also enhancing performance stability and reducing error rates across datasets. This work supports the potential of advanced AI techniques to improve early detection of skin cancer, aiding clinicians in making reliable, data-driven treatment decisions [27].

The remainder of this paper is structured as follows. Section two reviews the theoretical background, including foundational concepts of RNNs, DWT, and PCA. Section three outlines the proposed methodology, detailing the preprocessing pipeline, dataset description, feature extraction process, and the architecture of the RNN-based classification model. Section four presents the experimental results, performance evaluation metrics, and a comparative analysis with existing state-of-the-art methods. Finally, Section five provides concluding remarks and outlines future directions for research.

2. Background Review

2.1 Recurrent Neural Networks (RNNs)

Text, audio signals, and time series are examples of sequential data that can be handled by RNNs, a form of artificial neural networks. Their ability to memorize information from earlier phases in the sequence and utilize it to process or produce outputs in later steps is the fundamental concept underlying these networks. They are therefore ideal for activities requiring a grasp of context or reliance on data order [16].

The dependence relationships between data points are what RNNs rely on, as opposed to standard neural networks, which assume that all inputs are independent of one another. This is achieved by using a hidden state that retains information from the previous time step and combines it with the input at the current step to produce the output. The process of updating the hidden state can be expressed mathematically below.

$$h_t = f(W_h h_{t-1} + W_x x_t + b) \qquad (1)$$

Where: f: is the activation function, which can be a nonlinear function.

h_t: is the hidden state at time step t.

h_{t-1}: is the hidden state from the previous time step $t-1$.

x_t: is the input at time step t.

W_h: is the weight matrix for the hidden state from the previous time step.

W_x: is the weight matrix for the current input x_t.

b: is the bias term.

However, there are several difficulties in training RNNs, particularly the exploding gradient and vanishing gradient issues. The back propagation algorithm is used to calculate updates during training, which causes these issues and hinders the network's ability to learn long-term temporal patterns. RNNs have evolved into more advanced architectures such as the Gated Recurrent Unit (GRU) and Long Short-Term Memory (LSTM), to address these issues. These architectures incorporate techniques to manage the flow of information over time, making them more effective at managing sequential data.

2.1.1 Long Short-Term Memory (LSTM)

The vanishing gradient issue is addressed by Long Short-Term Memory (LSTM), which introduces memory cells and controls the information flow. It's consists of three gates as follow:

1. Input Gate: The amount of fresh data that is added to the memory cell is controlled by this gate.
2. Forget Gate: Manages the data that is erased from the memory cell.

3. Output Gate: For the current time step, the output gate regulates the amount of data that is taken out of the memory cell.

LSTM networks are effective at modeling long-term temporal dependencies, making them particularly suitable for applications such as machine translation and other tasks that require maintaining contextual information over extended sequences. LSTMs are computationally intensive and difficult to parallelize, making them slow to train and inefficient for long sequences due to the quantity of gates.

The architecture of the RNN used in this system consists of the following components:

- LSTM Layers: Long Short-Term Memory units are utilized to enhance feature processing by addressing the vanishing gradient issue and capturing patterns in the data. [20]
- Dropout Layers: Regularization is applied to improve generalization and avoid overfitting. [21]
- Batch Normalization: This stabilizes and accelerates training by normalizing activations. [22]

2.2 Discrete Wavelet Transform (DWT)

DWT is a potent tool for multi-resolution analysis that simultaneously captures frequency and spatial information. To enable multi-resolution analysis, this transformation depends on splitting the image into four sub-bands. LL sub-band represents the image's low-frequency components while maintaining its main characteristics and general structure. HL, LH and HH sub-bands capture high-frequency details in the following directions respectively: horizontal, vertical, and diagonal. The mathematical representation of the DWT is given as:

$$x_j[k] = \Sigma_{n=0}^{N-1} x[n].\psi_{jk}\ (n) \quad (2)$$

Where:

$x_j[k]$: represent the DWT coefficients at level j and scale k.

$x[n]$: is the original signal.

ψ_{jk} :is the wavelet function at level j and position k, which is derived from scaling and shifting operations on the base wavelet function.

N: is the number of samples in the original signal.

One of DWT key strengths is its ability to perform multi-resolution analysis, allowing for in-depth examination of signals while preserving critical information. Additionally, DWT is highly effective in data compression applications, significantly reducing data size without substantial loss of quality. However, DWT also presents certain limitations. It is sensitive to boundary conditions, which may result in distortions at the signal edges. Moreover, DWT can be computationally more intensive than alternative techniques such as the Fourier Transform, requiring greater processing power. Despite these challenges, DWT remains a powerful and widely adopted tool across scientific, engineering, and medical domains.

2.3 Principal Component Analysis (PCA)

One statistical technique for lowering the dimensionality of data while maintaining the greatest amount of variability is PCA to do this, the original data is converted into a new coordinate system with uncorrelated axes (called primary components) arranged according to how much variance they can capture. The mathematical representation of PCA is described as:

$$Z = \chi' V_k \quad (3)$$

Where:

Z**:** The transformed data (principal components).

χ': The mean-centered data matrix.

V_k: The matrix of the first k eigenvectors (principal components) [29].

The core steps in PCA include:

1. Mean Centering: To achieve zero mean for all features, subtract the feature means from the data:

$$x'_{ij} = x_{ij} - \mu_j \quad (4)$$

Where: x_{ij} is the original value, and μ_j is the mean of feature j.

2. Covariance Matrix Calculation: Compute the covariance matrix to understand the relationships between features:

$C = \frac{1}{n-1} X' \top X'$ (5)

Where C is the covariance matrix, and n is the number of samples.

3. Eigen Decomposition: Determine the covariance matrix's eigenvalues and eigenvectors:

$$Cv = \lambda v \quad (6)$$

Where: λ: The eigenvalues represent the variance explained by each principal component.

v: The eigenvectors define the directions of the principal components.

4. Dimensionality Reduction: To project the data into a lower-dimensional space, choose the top k eigenvectors that match the biggest eigenvalues.

PCA is especially useful in situations where high-dimensional data presents difficulties, such over fitting or computing inefficiency.

3. Skin Cancer Detection system

The design of the image recognition system generally involves the collection data, feature extraction/selection, model selection or training, and evaluation. This part describes the design of the detection system for skin cancer detection of digital images. The design process is divided into two main phases: the training phase and the testing phase.

- Training Phase: At the beginning of the training phase, a collection of digital images is used as training data. The data is then prepared by using the pre-processing phase. The image characteristics may then be extracted using a few features extraction techniques. The retrieved features can then be used to learn a selected model. The categorization model, which takes into account the phase's main goal, is the end product of this phase.
- Testing Phase: During this phase, the system's capacity to approximate system generalization and classify patterns into appropriate groups is tested using the new data.

3.1 Data Collection

The comprehensive ISIC dataset, available on the Kaggle platform, is utilized. This dataset contains high-quality medical images of various skin lesions, providing balanced representation across different classes. The ISIC dataset is a reliable and widely used resource in the field of skin cancer diagnosis using artificial intelligence. The final dataset used in this research consists of 120 images of benign lesions (non-threatening cases) and 1517 images of malignant lesions

(cases require urgent medical intervention). A larger number of malignant images were chosen to provide broader representation for this class during training, which helps improve the model's performance and accuracy in identifying malignant lesions. Figure 1 shows examples of benign and malignant lesions.

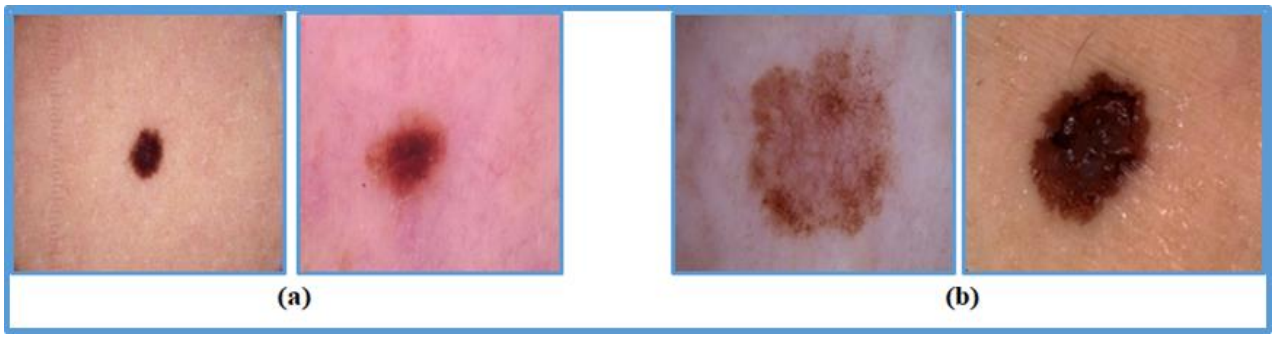

Figure 1: Sample images from the dataset, (**a**) benign (**b)** Malignant.

3.2 Pre-processing

The proposed system enhances input images for skin cancer detection and classification through a series of pre-processing steps. For consistent feature extraction, all images in this work are resized to 140 × 140 pixels to ensure uniformity in dimensions. To further standardize the input data and improve the model stability during training, the pixel values are normalized to a range between 0 and 1, where 0 represents black and 1 represents white. As color information may contribute to more accurate classification, the system retains the original color channels rather than converting the images to grayscale.

3.3 Features Extraction

Feature extraction is a crucial stage in the proposed skin cancer classification system. Its primary goal is to transform raw image data into meaningful and discriminative features for effective classification. The system utilizes a combination of DWT and PCA for both feature extraction and dimensionality reduction. The resulting features are then fed into RNN to enhance classification performance. DWT breaks the images down into four sub-bands (LL, LH, HL, HH). Only the LL sub-band is used in the proposed system as it eliminates high-frequency noise while preserving the most important characteristics. This effective representation lowers the computational complexity without losing the feature set's correctness. To further optimize the features extracted from the LL sub-band, PCA is applied. This technique reduces dimensionality by identifying and retaining the most significant components of the feature set, thereby ensuring an efficient representation while minimizing redundancy. The number of principal components is set to 100, striking a balance between classification accuracy and computational efficiency.

3.4 RNN Model for classification

This study addresses the classification of skin lesions as a two-class problem: benign and malignant. The problem is tackled using a deep learning approach with RNN-based architecture, leveraging LSTM layers for enhanced feature processing. RNN processes the reduced feature set for further classification and analysis. RNN are typically used for sequential data and captures complex dependencies within the feature set. Therefore, the system reshapes the extracted feature vectors into a two-dimensional sequential format to leverage the powerful processing capabilities

of RNNs, thereby enhancing the system's ability to differentiate between benign and malignant lesions. The architecture of RNN in this study consists of a sequence of LSTM layers with dropout and batch normalization applied after each layer. The model learns hierarchical patterns in the data by capturing spatial features through the LSTM layers.

Input: the input feature vectors, obtained from DWT and subsequently reduced in dimensionality using PCA, are reshaped into a two-dimensional sequential format suitable for processing by the LSTM layers.

LSTM Layers:

The core of the architecture comprises four LSTM layers:

- First LSTM Layer: Contains 128 units and outputs sequences for the subsequent layer. Dropout is set at 25%, and batch normalization is applied to stabilize activations.
- Second LSTM Layer: Composed of 64 units, this layer also outputs sequences. Dropout is set at 20%, and batch normalization is included.
- Third LSTM Layer: This layer has 32 units and outputs sequences to maintain temporal dependencies. A dropout rate of 20% is applied along with batch normalization.
- Final LSTM Layer: Includes 32 units and outputs a feature vector used for classification. Dropout is set to 25% to ensure effective regularization.

Output: is the binary classification via a sigmoid-activated neuron.

In this classification setup, the model is trained on labeled data (xi, yi), where $xi \in R^d$ represents the input features vector (image data), and yi is the binary label (0 for benign and 1 for malignant). The RNN model learns to optimize the binary cross-entropy loss function to classify the images accurately. The model is optimized using the Adam optimizer with binary cross-entropy loss to handle the binary classification problem. The training process is enhanced by call backs such as Early Stopping and ReduceLROnPlateau, ensuring efficient learning and avoiding over fitting. This approach offers a robust solution to the problem of skin cancer classification by combining the power of deep learning with specialized image pre-processing techniques, such as DWT and PCA, making it suitable for real-time medical applications. Figure 2 illustrates the block diagram of the proposed model for skin cancer detection.

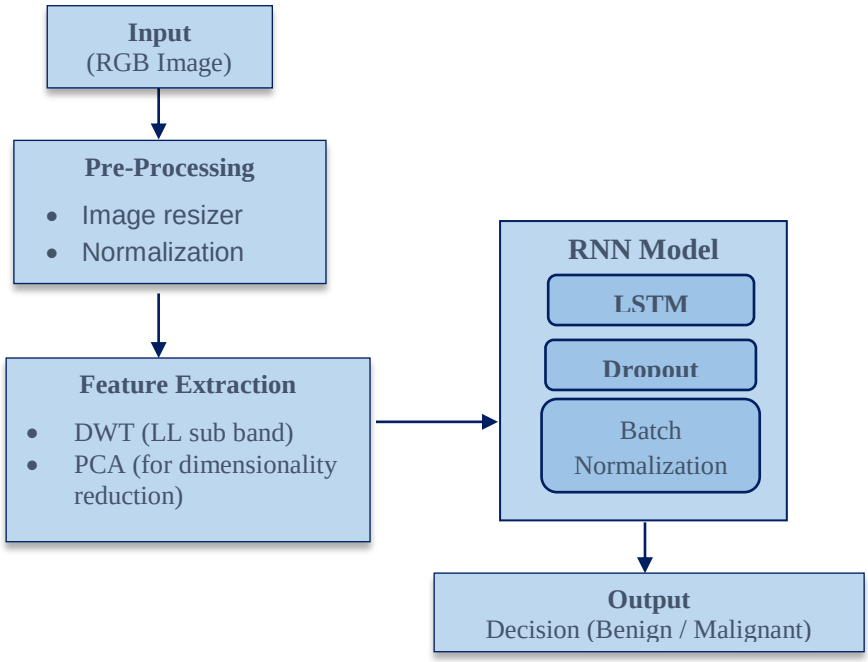

Figure 2: Illustrates how the model works.

4. Experimental Results

This study uses ISIC dataset [25] to evaluate the proposed model for skin cancer detection. All images are resized to 140 × 140 pixels to ensure uniformity in dimensions. Additionally, the pixel values are normalized to a range between 0 and 1, where 0 represents black and 1 represents white.

The performance of the proposed model is valuated in terms of accuracy, which is defined as the ratio of correctly classified samples to the total samples. Accuracy is calculated using the following formula:

$$Accuracy = \frac{(TP+TN)}{N} \quad (7)$$

$$N = TP + FP + TN + FN \quad (8)$$

Where: *TP* is the number of true positives, *TN* is the number of true negatives, *FP* is the number of false positives, FN is the number of false negatives, and *N* is the total number of samples**.**

Additionally, F1-score is utilized as metric in classification problem to measure the model's accuracy. This metric can be calculated as:

$$F1\_score = \frac{(Precision \cdot Recall)}{(Precision + Recall)}.2 \quad (9)$$

$$Precision = \frac{Tp}{(Tp+FP)} \quad , \quad Recall = \frac{Tp}{(Tp+FN)} \quad .$$

Where the used metrics (accuracy and F1-score) give a value between 0 and 1 where 1 is the optimal value.

The proposed model utilizes the RNN architecture with features that are extracted from LL sub-bands using DWT incorporating LSTM layers as a core component of RNN. The initial version of the model is evaluated using ISIC dataset which included two classes:

- Benign: 30 images.
- Malignant: 240 images.

This version shows promising results, achieving an accuracy of 92% in classifying skin cancer images.

To achieve better model performance, PCA is used to reduce the dimensionality of extracted features from LL sub-bands. PCA focuses on significant components, minimizing noise, and improving overall computational efficiency. Furthermore, a larger and more diverse dataset was utilized, comprising:

- Benign: 120 images.
- Malignant: 1517 images.

With these improvements, the model achieved a significantly higher accuracy of 98%, underscoring the effectiveness of combining advanced pre-processing techniques (DWT and PCA) with the RNN architecture. The unified model shows strong capability in differentiating benign from malignant skin lesions, underscoring the promise of deep learning in medical image classification.

Table 1: The performance of proposed Model

Feature extraction	**ISIC Dataset**	**Accuracy (%)**	**F1-score**
DWT	270 images (30 Benign and 240 Malignant)	94%	0.95
DWT+PCA	1637 images (120 Benign and 1517 Malignant)	98%	0.99

Figure 3 illustrates the accuracy progression of DWT with the RNN architecture during training and validation across different epochs, providing a clear view of the model's performance. The figure also shows the loss trends (measure of the model's error) highlighting the decrease in both training and validation loss over time. Additionally, Figure 4 illustrates the accuracy progression of DWT and PCA with the RNN architecture during training and validation across different epochs, providing a clear view of the model's performance. The figure also shows the loss trends (measure of the model's error) highlighting the decrease in both training and validation loss over time.

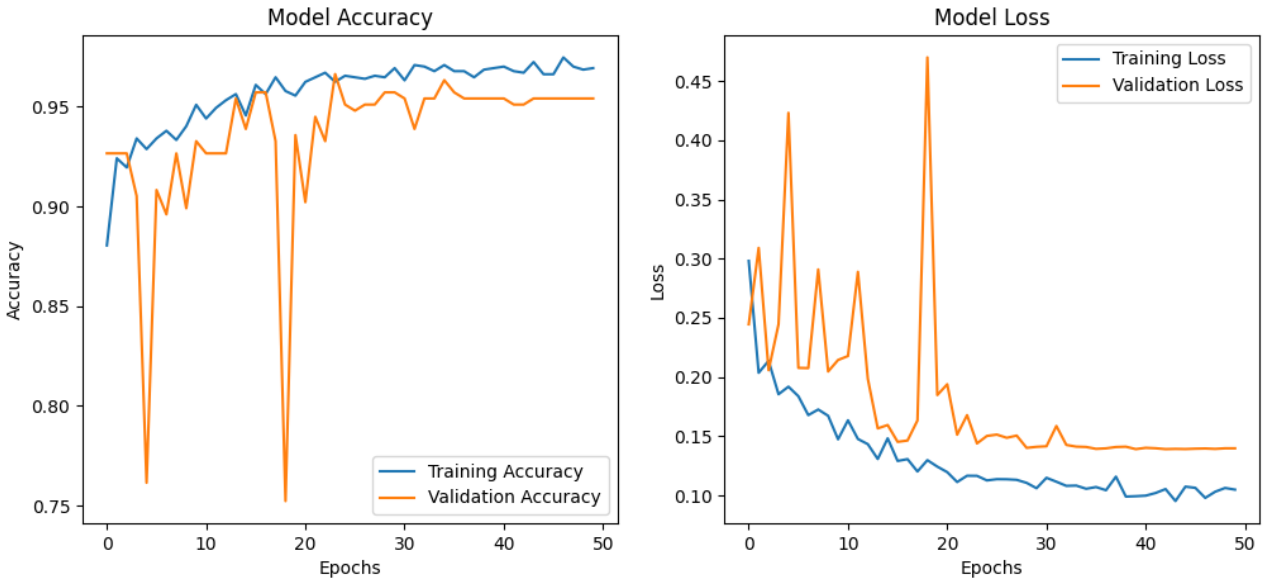

Figure 3: shows the accuracy and loss trends of the proposed model (DWT)

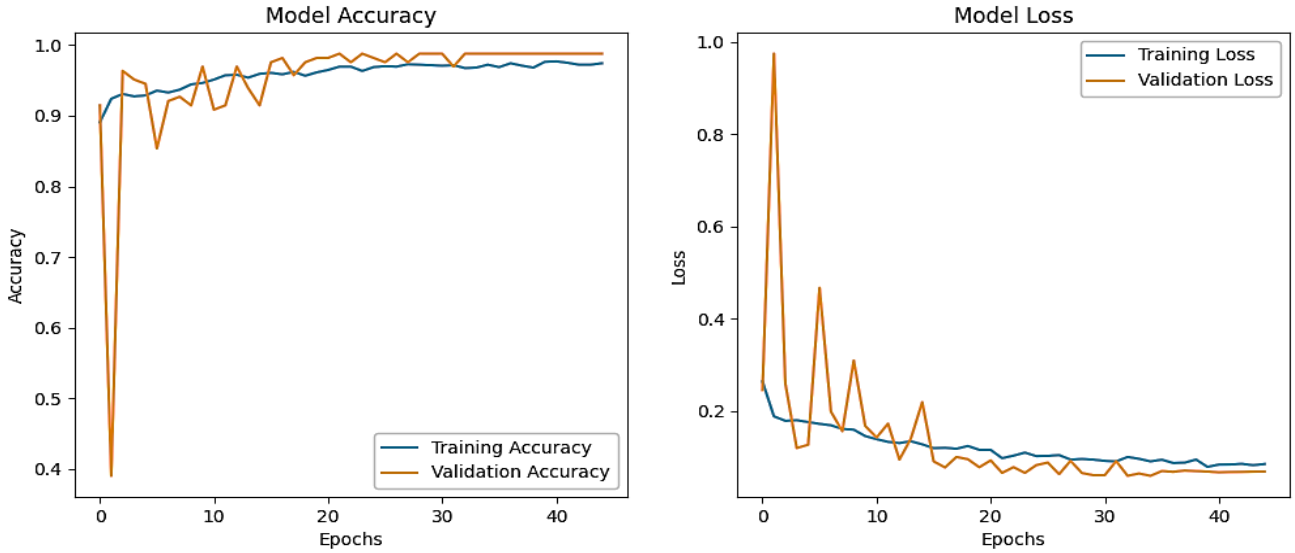

Figure 4: shows the accuracy and loss trends of the proposed model (DWT+PCA).

4.1 Cross-Validation (5-Fold Cross-Validation)

To ensure accurate and comprehensive evaluation of the proposed model, a 5-Fold Cross-Validation strategy was employed. This technique is widely adopted in machine learning research

as it mitigates the risk of bias associated with static data splits and provides a realistic assessment of the model's generalization ability.

In this study, the whole dataset was randomly divided into five equal subsets (folds). For each iteration, four folds were used to train the model while the remaining fold served as the test set. This process was repeated five times, with each fold acting as the test set exactly once. Upon completion, the mean and standard deviation of the performance metrics were calculated across the five runs, offering a more stable and reliable estimate of model performance. The results obtained from this validation process are summarized at Table 2.

Table 2: shows the summary of Cross-Validation Results

Metric	**Mean**	**± Standard Deviation**
Loss	0.1120	± 0.0273
Accuracy	96.33%	± 1.08%
F1-Score	0.9803	± 0.0057
AUC	0.9538	± 0.0320
False Positives (FP)	7.40	± 3.26
False Negatives (FN)	4.60	± 1.02

These results demonstrate a high degree of consistency in performance across folds, indicating the model's strong generalization capability and its robustness to variations in data splits. The high average accuracy is 96.33% and F1-score is 0.9803 that suggest a well-balanced trade-off between precision and recall. Moreover, the low rates of false positives and false negatives confirm the model's effectiveness in minimizing classification errors, which is especially crucial in medical applications like skin cancer detection.

Overall, this cross-validation strategy validates the effectiveness of combining DWT, PCA, and RNN architectures for the reliable and efficient classification of skin lesions, reinforcing the model's potential for real-world clinical deployment.

Figure 5 illustrates the training and validation performance of the proposed method using the RNN model. The left plot shows the accuracy progression across epochs, where the validation accuracy reaches a maximum of 96.33%. The right plot presents the loss trends, indicating a steady decrease in both training and validation loss over time, reflecting the model's ability to generalize effectively despite some fluctuations in validation loss. Figure 6 illustrates Final summary of performance metrics obtained from 5-Fold Cross-Validation. The bar chart displays the averaged values of AUC, F1-score, accuracy, and loss, as well as the average false positives (FP) and false negatives (FN). Figure 7 reflects the model's high classification accuracy and low error rates, confirming its robustness and suitability for reliable skin cancer detection. . Table 3 shows a comparison between the proposed approach and previous work in [12, 15].

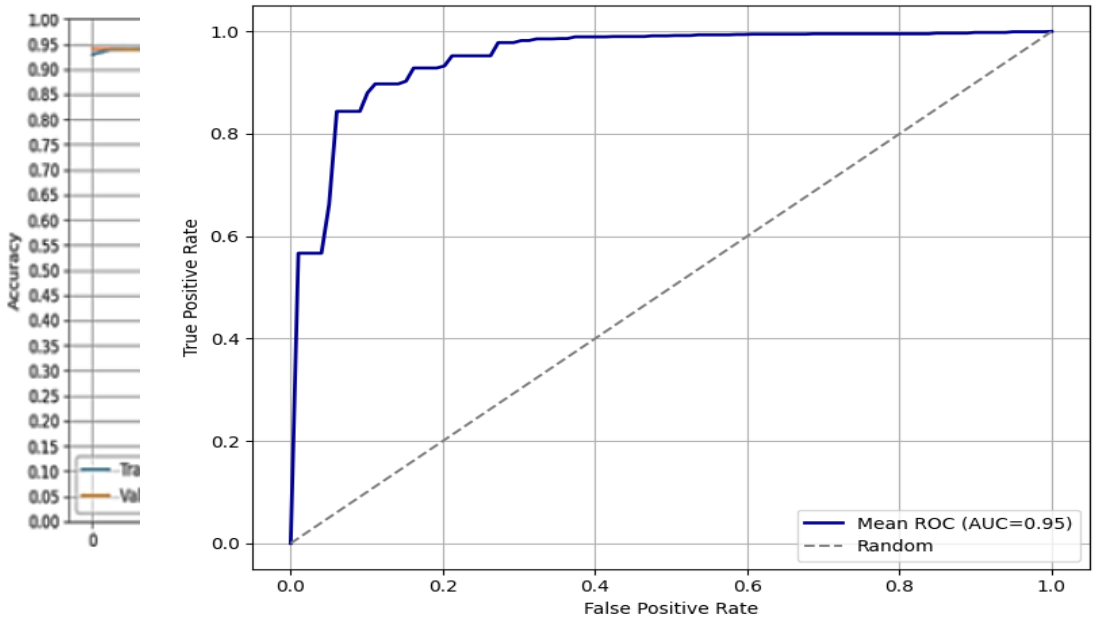

Figure 5. Accuracy and loss across 5-folds.

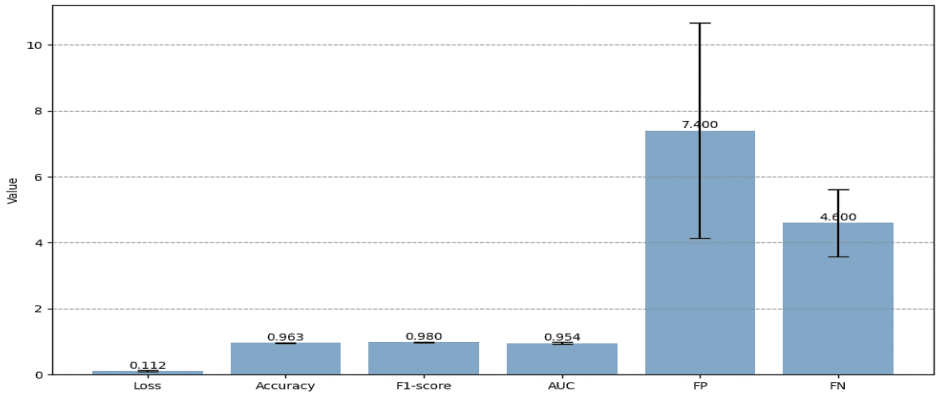

Figure 6. Final summary of performance

Figure 7. ROC curves for 5-Fold Cross-Validation.

Table 3 shows a comparison between the proposed model and the models in [12, 15]

Feature extraction	***Model***	***Accuracy***
DWT+PCA proposed method	RNN	98%
DWT+PCA+ 5-Fold Cross-Validation	RNN	96.33%
CNN in [15]	RNN	94.48%
DWT in [12]	CNN	94%

5. Conclusion and Future Work

This paper presented a deep learning framework for skin lesion classification, employing an RNN with LSTM layers. Features were extracted from the LL sub-bands of the DWT, followed by dimensionality reduction using PCA to retain the most informative features while enhancing computational efficiency. The proposed approach achieved a high classification accuracy of 98%. Furthermore, a 5-fold cross-validation was conducted, resulting in an average accuracy of 96.33%, confirming the robustness and generalizability of the model across different subsets of the dataset.

Future research will focus on expanding the dataset in both size and diversity to further improve model generalization. Additionally, integrating explainable AI techniques could provide clinicians with interpretable insights, supporting practical adoption in dermatological diagnostics.

References:

[1] American Cancer Society. (2023). *Skin Cancer*.

[2] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). *Dermatologist-level classification of skin cancer with deep neural networks*. Nature, 542(7639), 115–118.

[3] Mitchell, T. M. (1997). Machine Learning. McGraw-Hill Education.

[4] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

[5] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by backpropagating errors. Nature.

[6] Mallat, S. (1999). A Wavelet Tour of Signal Processing: The Sparse Way. Academic Press.

[3] Esteva, A., et al. (2017). Skin Cancer Detection using Deep Learning. IEEE Conference Publication.

[4] El-Khatib, H., Popescu, D., & Ichim, L. (2020). Deep Learning–Based Methods for Automatic Diagnosis of Skin Lesions. Journal of Medical Imaging, vol. 27, no. 4, pp. 123-135.

[5] Kim, C.-I., Hwang, S.-M., Park, E.-B., Won, C.-H., & Lee, J.-H. (2021). Computer-Aided Diagnosis Algorithm for Classification of Malignant Melanoma Using Deep Neural Networks. Sensors, 21(16), 5551. https://doi.org/10.3390/s21165551

[6] Jain, S., Singhania, U., Tripathy, B., Nasr, E. A., Aboudaif, M. K., & Kamrani, A. K. (2021). Deep Learning-Based Transfer Learning for Classification of Skin Cancer. Sensors, 21(23), 8142. https://doi.org/10.3390/s21238142

[7] Hussain, S. I., et al. (2022). Melanoma Identification Through X-ray Modality Using Inception-v3 Based Convolutional Neural Network. IEEE Transactions on Medical Imaging, vol. 41, no. 5, pp. 1234–1242.

[8] Hussain, S. I., et al. (2024). Skin Cancer Detection and Classification Using Neural Network Algorithms. MDPI Sensors, vol. 24, no. 7, pp. 567–579.

[9] Bhatti, Z. I., et al. (2023). Skin Cancer Detection and Classification Using Deep Neural Networks. Springer Lecture Notes in Computer Science, vol. 13692, pp. 89–101.

[10] Ansari, M., et al. (2023). Skin Cancer Detection Using Deep Neural Networks Based on ResNet and Inception Architectures. MDPI Applied Sciences, vol. 13, no. 15, pp. 3561.

[11] Tembhurne, J. V., Hebbar, N., Patil, H. Y., & Diwan, T. (2023). Skin cancer detection using ensemble of machine learning and deep learning techniques. Multimedia Tools and Applications, 82, 27501–27524.

[12] Claret, S. P. A., Dharmian, J. P., & Manokar, A. M. (2024). Artificial intelligence-driven enhanced skin cancer diagnosis: leveraging convolutional neural networks with discrete wavelet transformation. Egyptian Journal of Medical Human Genetics, 25, Article number: 50.

[13] Azizi, S., Biyabani, S. R., Yan, P., Tahmasbi, A., Kwak, J. T., Xu, S., Turkbey, B., Choyke, P. L., Pinto, P. A., Wood, B. J., Mousavi, P., & Abolmaesumi, P. (2018). Deep Recurrent Neural Networks for Prostate Cancer Detection: Analysis of Temporal Enhanced Ultrasound. PubMed Central.
[14] Ganguli, R., Lad, R., Lin, A., & Yu, X. (2023). Novel Generative Recurrent Neural Network Framework to Produce Accurate, Applicable, and Deidentified Synthetic Medical Data for Patients With Metastatic Cancer. JCO Clinical Cancer Informatics. PubMed Central.
[15] Zareen, S. S., Sun, G., Kundi, M., Qadri, S. F., Qadri, S., 2024. Enhancing Skin Cancer Diagnosis with Deep Learning: A Hybrid CNN-RNN Approach. Computers, Materials & Continua 74, pp. 1245-1258.
[16] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
[17] Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning Phrase Representations using RNN Encoder-Decoder. Empirical Methods in Natural Language Processing (EMNLP), 1724-1734.
[18] Olah, C. (2015). Understanding LSTM Networks. Colah's Blog. Retrieved from http://colah.github.io/posts/2015-08-Understanding-LSTMs/
[19] Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation. MIT Press.
[20] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Journal of Machine Learning Research. MIT Press.
[21] Ioffe, S., & Szegedy, C. (2015). Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. arXiv preprint. arXiv.
[22] Mallat, S. (1989). A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. IEEE Transactions on Pattern Analysis and Machine Intelligence. IEEE.
[23] Almarimi, A. F., & Salem, A. M. (2025). Machine Learning using Simple Linear Regression. Bani Waleed University Journal of Humanities and Applied Sciences, 10(3), 178-184.
[24] Jolliffe, I. T. (2002). Principal Component Analysis. Springer Series in Statistics. Springer.
[25] Shlens, J. (2014). A Tutorial on Principal Component Analysis. arXiv preprint arXiv:1404.1100. Retrieved from https://arxiv.org/abs/1404.1100.
[26] Codella, N., Rotemberg, V., Tschandl, P., et al. (2019). The ISIC Skin Image Analysis Collaboration: 2018–2019 Challenge Datasets and Techniques. arXiv preprint arXiv:1902.03368. Retrieved from https://arxiv.org/abs/1902.03368.
[27] Soglia, S. (2025). Artificial Intelligence-Based Integrated Approach for Skin Cancer Recognition. PhD thesis, University of Brescia.
[28] Hosseinzadeh, M., Hussain, D., Mahmood, F. M. Z., Alenizi, F. A., Noroozi Varzeghani, A., Asghari, P., Darwesh, A., Malik, M. H., & Lee, S. W. (2025). A model for skin cancer using combination of ensemble learning and deep learning. PLOS ONE, 20(4), e0301275. https://doi.org/10.1371/journal.pone.0301275