SvedbergOpen
DISSEMINATION OF KNOWLEDGE
International Journal of Artificial Intelligence and Machine Learning
Publisher's Home Page: https://www.svedbergopen.com/
Research Paper Open Access

IRT-Simulated Hyperspectral Deep Learning For Accurate Melanoma Detection

P Babu[1*], Dr. T. Meyyappan[2]

[1*]PhD Resesearch Scholar, Department of Computer Science, Alagappa University Karaikudi, Tamilnadu, India
meyyappant@alagappauniversity.ac.in
[2]Senior Professor, School of computational science, Alagappa University Karaikudi, Tamilnadu, India
pb.darsana.alu@gmail.com

Abstract
Early detection of melanoma is crucial for improving survival rates, yet current diagnostic methods often rely on subjective analysis and limited imaging modalities. This study proposes a novel IRT-simulated hyperspectral deep learning framework for skin cancer classification using the Kaggle Skin Cancer: Malignant vs. Benign dataset. Initially, RGB dermoscopic images are synthetically transformed into simulated Infrared Thermography (IRT) thermal maps, generating steady-state and rewarming phases to reveal subsurface lesion characteristics. These are further converted into hyperspectral-like representations through spectral decomposition, producing rich spectral information for lesion differentiation. Pre-processing involves background removal, morphological hair filtering, and hue-gamma correction to enhance lesion visibility. Spectral-spatial features are then extracted and refined using the Whale Optimization Algorithm (WOA) to retain the most relevant and non-redundant features. Classification is performed using NASceptionNet, a hybrid model combining NASNet and XceptionNet, enabling robust and accurate lesion classification. The framework is evaluated using multiple performance metrics, demonstrating its potential as a non-invasive, accurate, and efficient screening tool for early melanoma detection, bridging the gap between dermoscopic imaging and advanced spectral analysis. The proposed model demonstrated remarkable improvement after IRT conversion. Before IRT conversion, the model achieved 98.04% accuracy, 97.88% precision, 98.21% recall, and 98.04% F1-score, with an MCC of 96.09. After IRT conversion, performance increased to 99.12% accuracy, 99.05% precision, 99.18% recall, and 99.11% F1-score, with an MCC of 98.23. These results indicate a consistent 1–2% performance boost across all metrics, confirming the effectiveness of IRT conversion in enhancing diagnostic accuracy.

Keywords: Skin Cancer, Deep Learning, IRT Conversion, Classification, Accuracy, Feature Optimization, Explainability.

1. Introduction

Melanoma is one of the most malignant skin cancers, which has a rapid development and a high mortality rate in case it is not diagnosed at an early stage. Conventional methods of diagnosis, such as dermoscopic and histopathology, are highly dependent on the experience of dermatologists and thus exhibit inconsistency in diagnosis because of subjective analysis. The increasing worldwide prevalence of melanoma has emphasized the necessity of more precise, dependable, and accessible diagnostic tools to assist with early identification and prompt treatment. Recent progress in artificial intelligence (AI), especially deep learning, has demonstrated potential in medical image analysis. CNNs and other deep learning architectures have the ability to automatically learn discriminative features in dermoscopic images, which outperform conventional machine learning techniques that require handcrafted features. Not only do these models decrease diagnostic subjectivity but also show human-comparable and in some cases better accuracy in classifying melanoma against benign skin lesions. The development of such progress highlights the revolutionary potential of deep learning in the field of computer-aided dermatological diagnosis. The potential of deep learning in melanoma detection systems is the possibility to enhance clinical decision-making, large-scale screening, and accessibility in resource-limited settings. Nonetheless, data imbalance, variability of image quality, and explainable AI are still important research

directions. By solving these problems with sophisticated architectures, hybrid solutions, and thorough validation, it is possible to open the path to reliable real-world adoption. Therefore, deep learning-based melanoma detection is one of the most promising technologies in intelligent healthcare that can help to eliminate diagnostic errors and eventually increase patient outcomes.

Although there have been great advances, current deep learning models in melanoma detection have their share of limitations. The limited annotated datasets may lead to overfitting of models like InceptionV3 and ResNet50, decreasing their applicability to the real clinical environment. VGG16 has a high computational cost and is therefore not ideal when it comes to implementation in resource-limited settings. Moreover, DenseNet121 and EfficientNet have a problem with class imbalance, as benign lesions are much more than melanoma cases, which causes biased predictions. The inability to interpret in most models is another issue that is likely to cause clinicians to not completely trust AI-based diagnostic suggestions. Melanoma is a major health concern in the world because of its fast development rate and high mortality rate when not diagnosed early. Although dermoscopic is a well-established method of clinical diagnosis, it is highly subjective to physician experience and restricted by the 2D quality of RGB images, which are unable to reflect subsurface lesion features. Conventional deep learning models, albeit efficient in feature extraction, are mainly based on visible spectrum images and do not consider important thermal and spectral information related to melanoma development. Infrared thermography (IRT) has demonstrated potential in the detection of metabolic changes in cancerous tissue, yet stand-alone thermal images lack the spatial and spectral complexity required to make accurate classification. Likewise, hyperspectral imaging can offer detailed spectral signatures to differentiate lesions, but it is expensive, time-consuming, and not frequently accessible in everyday clinical practice.

The gaps in such an approach require a new framework that combines the availability of dermoscopic imaging and the diagnostic power of hyperspectral and thermal imaging. This paper attempts to simulate IRT and hyperspectral data on standard dermoscopic images to create rich spectral-spatial features without the costly equipment. This would allow better characterization of the lesions, better diagnostic performance, and a scalable, non-invasive method of early melanoma screening.

1.1 Key contribution

Following are the key contribution of the study,

- Initially, the Skin Cancer: Malignant vs. Benign dataset from Kaggle was collected, and RGB dermoscopic images were prepared for transformation into simulated Infrared Thermography (IRT) form.
- The RGB images were converted into IRT-style thermal maps to mimic real thermal scans. Steady-state and rewarming phase simulations were generated to highlight subsurface lesion patterns.
- The simulated thermal images were transformed into hyperspectral-style bands through spectral decomposition, enriching the data with fine spectral details to improve malignant–benign differentiation.
- Pre-processing was applied using background removal to isolate lesion regions, morphological hair filtering to eliminate hair artifacts, and hue-gamma correction to enhance lesion contrast.
- Spectral-spatial features were extracted from the pre-processed hyperspectral data, and the Whale Optimization Algorithm (WOA) was employed to select the most informative and non-redundant features while balancing exploration and exploitation.
- The optimized features were classified using NASceptionNet, a hybrid CNN combining NASNet and XceptionNet to capture complex lesion patterns, improving classification accuracy and generalization.
- The proposed model's performance was evaluated using accuracy, precision, recall, specificity, F1-score, NPV, MCC, FPR, and FNR, demonstrating its capability as a reliable non-invasive melanoma detection framework.

The paper is organized as follows: Section 1 introduces the research background and motivation. Section 2 reviews related work on melanoma detection. Section 3 details the proposed IRT-simulated hyperspectral deep learning framework. Section 4 presents experimental setup and results. Section 5 discusses overall conclusion of the study.

2. Literature review

Following are the recent literatures related to this study,

Mahmud et al. (2025) created explainable deep learning models to improve the early detection of melanoma with dermoscopic images. Their method combined Global Average Pooling, Batch Normalization, and the use of advanced activation functions such as ReLU and Swish to increase the accuracy on a variety of datasets. The model achieved an accuracy of more than 95 percent and had good

generalization capability. Notably, explainable AI methods like Grad-CAM and saliency maps gave interpretability, which gave clinicians a clear picture of the decision-making process. This not only makes the model accurate but also reliable, which is the significant obstacle of clinical adoption in AI-based dermatology diagnostics.

Swapno et al. (2025) suggested a hybrid pipeline of EfficientNetV2L and LightGBM to extract deep features and perform efficient classification, respectively. With a dataset of 3,297 images and data augmentation, their ensemble performed very well with a validation accuracy of 99.93% and ROC-AUC scores of 0.98. The strong precision, recall and F1-scores on both benign and malignant cases highlighted the strength of the model. Their analysis also showed the advantages of gradient boosting integration, which guarantees computational efficiency without compromising accuracy. This framework offers a scalable clinical skin cancer detection solution.

chawla et al. (2025) suggested a hybrid CNN model that incorporated VGG-16 and EfficientNet to classify dermoscopic images. They achieved an accuracy of 96.3%, recall and precision of 94.4 and 95.7, respectively. The research demonstrated the potential of hybrid deep learning to enhance the accuracy of diagnosis by combining the respective advantages of the two architectures. The model minimizes the use of expensive pathological testing, providing an economical diagnostic tool. Their study highlights the importance of AI in early diagnosis, increased accessibility and efficiency of skin cancer care.

Kaur et al. (2025) proposed a hybrid CAD system to detect melanoma which consisted of a combination of preprocessing, segmentation, and classification. They used ISIC 2020 data and performed morphological operations to reduce noise and subsequently segmented and classified using deep networks. The accuracy of the segmented-cleaned images was higher at 93.40 % than that of raw images classification. In their work, they highlighted the importance of preprocessing and segmentation in enhancing CAD systems. The framework has potential in clinical use where speed and accuracy are important as the test time per image is reduced.

Akinrinade and Du (2025) discussing the issue of healthcare accessibility, especially in rural areas, and deep learning models to detect skin cancer. They focused on addressing the problem of dataset imbalance by using transfer learning, augmentation, and GAN-based synthesis. They showed that they could achieve robust classification results on small data by experimenting with various deep learning backbones, ensemble strategies, and sampling strategies. This study demonstrated the possibility of using few-shot learning to facilitate accurate diagnosis in settings where resources are limited. Their input is an indication of the social significance of AI in addressing the disparity between underserved communities and complex cancer diagnostics.

Patil (2025) proposed a hybrid Transformer-CNN model that can combine local feature extraction and global contextual learning. Transformers were added to CNNs to learn long-range dependencies in dermoscopic images. The hybrid model had better robustness and precision in various types of skin cancer such as melanoma, basal cell carcinoma, and squamous cell carcinoma. Their study supported the significance of transformer integration in medical imaging, which is still in its infancy in terms of adoption. This method enhances the detection of skin cancer, since it offers both high-resolution and low-resolution information of the lesion and image context respectively.

Mohamed et al. (2025) suggested a multi-modal hybrid model (MiSC) which combined dermoscopic images with clinical metadata to classify skin cancer. Their model used MobileNetV2 to extract image-based features, and conventional ML models such as logistic regression and random forests to analyze the metadata. Applied to the PAD-UFES-20 dataset, the method demonstrated 95.6 percent accuracy, which indicates the diagnostic value of clinical metadata. They concluded that non-image data can be used to dramatically enhance classification performance over image-only methods. This multi-modal approach is representative of the real-world clinical scenario in which imaging is augmented with patient data.

Dorathi Jayaseeli et al. (2025) introduced the DSC-EDLMGWO framework, which combines SE-DenseNet as a feature extraction tool with an ensemble of deep learning models, which are optimized using the Gray Wolf Optimization (GWO) algorithm. They used CLAHE to enhance contrast and Wiener filter to remove noise. The framework recorded high accuracy of 98.38% and 98.17% on HAM10000 and ISIC datasets respectively. The model used ensemble deep learning combined with metaheuristic optimization to solve both the feature robustness and hyperparameter tuning problems. Their study shows a high-performance diagnostic system that has a high level of performance compared to conventional models.

Table 1: Research gap

Author(s), Year	Proposed Technique	Limitations	Significance	Performance (Scores & Metrics)

Mahmud et al. (2025) [13]	Explainable DL with GAP, BN, Dropout, ReLU & Swish + Grad-CAM, Saliency Maps	Dataset variability; may face overfitting with small/imbalanced datasets	Achieved robust accuracy; provided interpretability for clinical trust	Accuracy: 95.23%, 96.48%; explainable visualization tools validated predictions
Swapno et al. (2025) [14]	EfficientNetV2L for feature extraction + LightGBM classifier	Requires high-quality datasets; may need large-scale validation	Hybrid ensemble boosted accuracy & computational efficiency	Training Acc: 99.57%; Validation Acc: 99.93%; Test Acc (5th fold): 99.90%; Precision: 0.98–0.99; Recall: 0.98; F1-score: 0.98; ROC-AUC: 0.98
Chawla et al. (2025) [15]	Hybrid CNN combining VGG-16 + EfficientNet	High computational cost; limited dataset diversity	Hybrid model outperformed standalone CNNs; reduced diagnostic error	Accuracy: 96.3%; Recall: 94.4%; Precision: 95.7%
Kaur et al. (2025) [16]	Hybrid CAD pipeline: preprocessing + segmentation (morphological ops + deep nets) + classification	Lower accuracy with raw images; preprocessing adds complexity	Demonstrated impact of preprocessing/segmentation on CAD reliability	Accuracy: 93.40% with preprocessed images; Test time: 1.3s per image
Akinrinade & Du (2025) [17]	DL models addressing imbalance (transfer learning, augmentation, GANs, ensembles)	Performance depends on synthetic data quality; limited real-world trials	Proposed scalable, cost-effective system for underserved areas	High accuracy across DL backbones; precise metrics not reported
Patil (2025) [18]	Hybrid Transformer-CNN model	Transformer integration increases complexity & computational needs	Captured both local (CNN) & global (Transformer) features	Improved accuracy & robustness across melanoma, BCC, SCC (quantitative metrics not specified)
Mohamed et al. (2025) [19]	MiSC: Multi-modal (dermoscopic images + metadata) using MobileNetV2 + logistic regression + random forest	Requires metadata availability; may limit deployment in image-only scenarios	Showed added diagnostic value of patient metadata	Accuracy: 95.6%; Precision: 96.8%; Recall: 95.6%; F1-score: 95.7%
Dorathi Jayaseeli et al. (2025) [20]	DSC-EDLMGWO: SE-DenseNet + ensemble DL (LSTM, ELM, SSDA) + Gray Wolf Optimization	Complex pipeline; high computational requirements	Superior accuracy with ensemble + metaheuristic optimization	Accuracy: 98.38% (HAM10000), 98.17% (ISIC); Outperformed baseline models

2.1 Problem statement

Melanoma is one of the most aggressive and life-threatening forms of skin cancer, with patient survival highly dependent on early and accurate detection. Traditional diagnostic techniques, such as dermoscopic examination and histopathology, remain effective but are time-consuming, subjective, and reliant on specialist expertise. Automated deep learning approaches have emerged to assist diagnosis; however, most depend exclusively on RGB dermoscopic images, limiting their capacity to detect deeper tissue structures and subtle subsurface variations. Infrared Thermography (IRT) and Hyperspectral Imaging (HSI) have shown potential in revealing temperature variations and spectral signatures linked to malignant lesions, but the scarcity of publicly available IRT/HSI datasets hinders robust model training and validation. Existing deep learning models also face several challenges: inadequate spectral-spatial exploitation in conventional CNNs reduces classification accuracy, small or imbalanced datasets increase overfitting risks, and the absence of hybrid optimization strategies results in suboptimal feature selection and network

tuning. Furthermore, many current solutions are computationally expensive, restricting real-time applicability. Therefore, there is a pressing need for a robust, hybrid deep learning framework capable of simulating IRT/HSI data from accessible datasets, extracting optimal spectral-spatial features, and employing efficient optimization strategies for accurate, fast, and non-invasive melanoma detection suitable for clinical use.

2.2 Research Motivation

The motivation behind this work is to overcome the limitations of existing melanoma detection models by introducing a framework that can simulate Infrared Thermography (IRT) and Hyperspectral Imaging (HSI) conditions from readily available RGB dermoscopic datasets. This approach addresses the scarcity of public IRT/HSI datasets while enabling the extraction of rich spectral-spatial information crucial for early and accurate diagnosis. By integrating the Whale Optimization Algorithm (WOA) for optimal feature selection and employing NASceptionNet, a hybrid of NASNet and XceptionNet, the model is designed to enhance classification accuracy, reduce false detections, and improve generalization. This non-invasive, cost-effective, and computationally efficient solution has the potential to assist dermatologists in real-time clinical decision-making, ultimately improving patient survival rates and reducing the global healthcare burden caused by melanoma.

3. Proposed model

The proposed workflow in Figure 1 integrates simulated Infrared Thermography (IRT) with hyperspectral transformation and advanced deep learning for accurate melanoma detection. Initially, RGB dermoscopic images from the Kaggle Skin Cancer: Malignant vs. Benign dataset is transformed into simulated IRT images, including steady-state and rewarming phases, to capture subsurface lesion patterns. These are further decomposed into hyperspectral-like bands through spectral decomposition, enriching spectral-spatial details essential for distinguishing malignant from benign lesions.

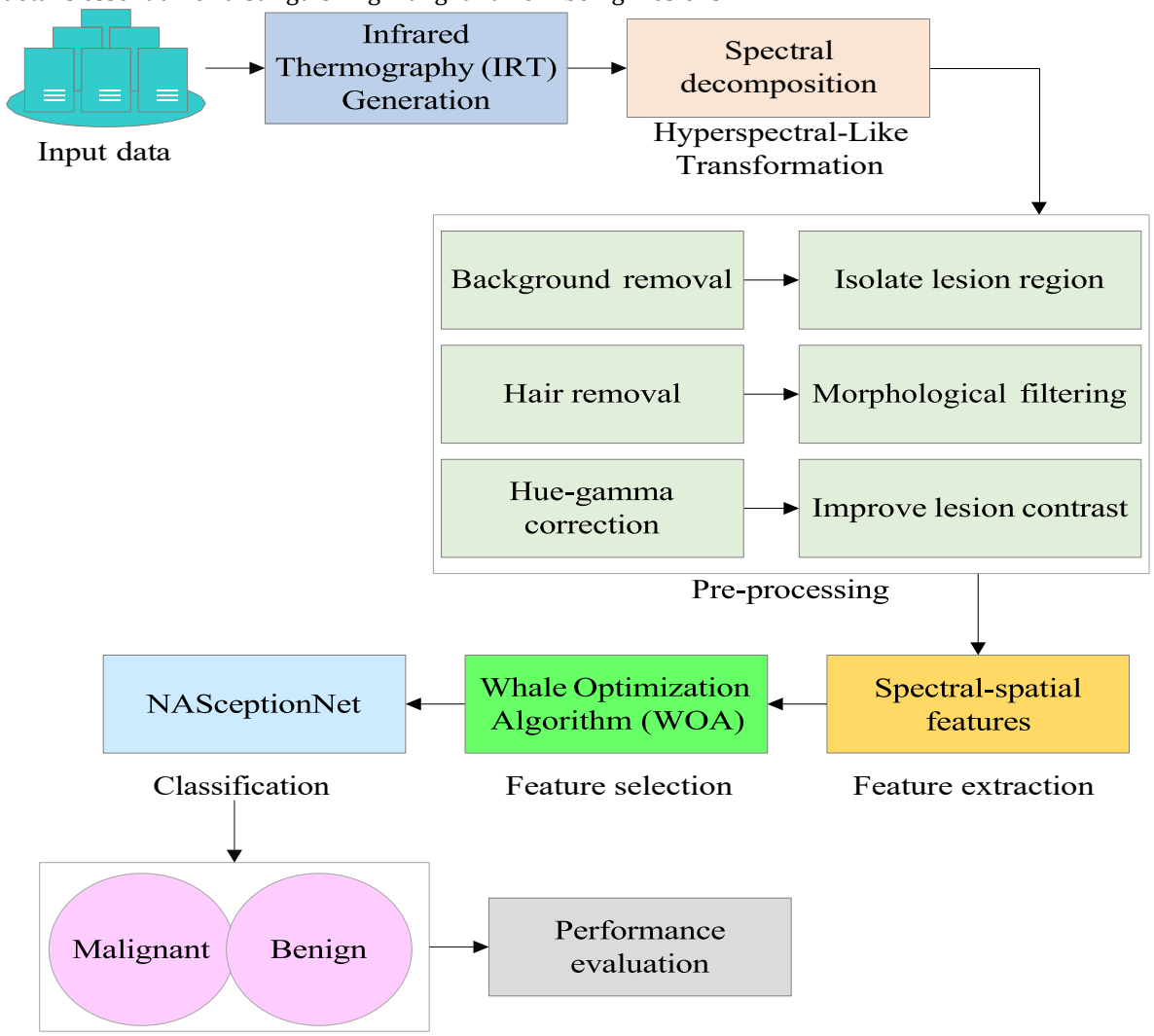


Figure 1: Overall architecture of the proposed model

Pre-processing is applied to remove background noise, hair, and contrast inconsistencies, ensuring cleaner lesion visibility. Spectral-spatial features are then extracted, capturing both textural and spectral characteristics. The Whale Optimization Algorithm (WOA) is employed for feature selection, balancing exploration and exploitation to retain the most relevant and non-redundant features. The optimized features are classified using NASceptionNet, a hybrid deep learning architecture combining NASNet's efficient architecture search with XceptionNet's depthwise separable convolutions, enabling detailed feature learning and improved generalization. Classification outputs are refined through thresholding to reduce false positives and negatives. Finally, performance evaluation metrics such as accuracy, precision, recall, specificity, and F1-score are used to validate the model. This integrated approach offers a robust, non-invasive, and clinically applicable melanoma detection framework with potential for real-world healthcare deployment.

3.1 Data collection

The Skin Cancer: Malignant vs. Benign dataset on Kaggle (by fanconic) is a balanced set of dermoscopic images obtained through the ISIC Archive, with two main folders, benign and malignant, containing around 1,800 images each to be used both as training and testing sets. The dataset includes RGB dermoscopic images in a 224 x 224-pixel format that can be easily used in deep learning pipelines. The division into train and test sets allows easy model building and testing. This dataset is sufficiently large, of high quality, and diverse, which makes it suitable to be used in binary classification tasks in skin lesion diagnosis. Its clear structure enables researchers to preprocess, train and test classifiers efficiently with confidence in the representativeness of both benign and malignant cases. Moreover, the following data has been collected using the following link, https://www.kaggle.com/datasets/fanconic/skin-cancer-malignant-vs-benign.

3.2 Infrared Thermography (IRT) Generation

The generation of infrared thermography (IRT) offers an alternative method of simulation of skin tissue thermal distribution based on dermoscopic RGB images that is non-invasive. Malignant lesions are more metabolically active and abnormally vascularized, which makes them produce more heat than benign lesions. To reproduce this effect, RGB dermoscopic images are artificially converted into thermal maps that simulate both the steady-state heat conduction and the rewarming process following a cooling stimulus. The steady-state phase records steady thermal gradients, whereas the rewarming phase records abnormal recovery patterns associated with tumor angiogenesis. These simulations give subsurface diagnostic information that is not seen in normal dermoscopic images. Synthetic IRT generation improves lesion characterization, and thus more discriminative features can be extracted by applying heat transfer equations to pixel intensity distributions. This method fills the gap between visual dermoscopic and physiological heat signatures, and enhances sensitivity in identifying malignancy when combined with machine learning classifiers.

Steady-State Heat Conduction is given in eqn. (1),

$$\nabla.k\Delta T+Q=0 \quad (1)$$

where k is the thermal conductivity, T defines the temperature and Q defines the metabolic heat source. Pennes' Bioheat Equation (Rewarming Phase) has been given in eqn. (2),

$$\rho c\partial T\partial t=\nabla.k\Delta T+\omega bcbTa-T+Qm \quad (2)$$

where ρ defines the tissue density, c is the specific heat, ωb representing the blood perfusion rate, cb is the blood specific heat, Ta defines the arterial blood temperature, Qm representing the metabolic heat.

3.3 Spectral Decomposition and Hyperspectral-Like Transformation

Based on the synthetic IRT maps (temperature field) T(x,y,t), we proceed to generate hyperspectral-like image cubes by mapping temperature to wavelength-dependent radiance and then decomposing those radiances into a set of spectral bands using the temperature as the wavelength. The abnormal thermal signature of malignant tissue generates unique spectral radiance curves when analyzed in terms of thermal emissivity and blackbody radiation models. A pseudo-hyperspectral cube is obtained by sampling the emitted radiance over a discrete set of wavelengths H(x,y,λi). The resulting product is produced which encodes spatial structure and wavelength-dependent thermal information. Dimensionality-reduction and unmixing algorithms (PCA, SVD, NMF, or linear spectral unmixing) are then used to isolate the major spectral signatures (endmembers) and their spatial abundances. These spectral-spatial representations augment discriminative features-e.g., variations in recovery dynamics, perfusion-related spectral slopes, and localized thermal heterogeneity-and can thus provide more input to downstream feature extraction and classification than RGB alone.

Spectral radiance (discrete wavelengths) using Planck's law with emissivity ε(λ) and it can be mathematically expressed using eqn. (3)-(4),

$$B\lambda,T=2hc2\lambda51ehc\lambda kBT-1 \quad (3)$$

$$L(x,y,\lambda i)=\varepsilon\lambda iB(\lambda i,Tx,y) \quad (4)$$

where h is the Planck constant, c defines the speed of light, kB is the Boltzmann constant, and λi are discrete band centers.

Form hyperspectral-like cube using the following eqn. (5),

$$Hx,y,\lambda i=Lx,y,\lambda i+\eta(x,y,\lambda i) \quad (5)$$

with η modelling sensor/noise and simulation errors.

Linear mixing / spectral unmixing (endmember model) is mathematically expressed using eqn. (6),

$$Hx,y,\lambda \approx r=1Rar(x,y)sr(\lambda) \quad (6)$$

or matrix form H≈A S, where sr(λ) are endmember spectra and ar(x,y) their abundance maps.

Dimensionality reduction (PCA / SVD) is mathematically expressed using eqn. (7),

$$\text{Reshape } H\longrightarrow Hp\times b;H=U\Sigma VT \quad (7)$$

retain top k components Hk=UkΣkVkT. The resulting outputs are either (a) down-sampled spectral-spatial feature maps (principal components / abundance maps) or (b) the entire simulated cube H(x,y,λi). These are fed directly to the spectral-spatial feature extraction module that calculates texture, spectral gradients, and statistical descriptors in order to select the WOA-based features and perform the NASceptionNet classification.

3.4 Pre-Processing

After the creation of hyperspectral-like IRT maps, a specific preprocessing step is used to enhance data quality and highlight the lesion-specific features prior to feature extraction. Background removal is done to isolate the lesion to the surrounding healthy skin so that classification is done on clinically relevant areas only. This segmentation is done by adaptive thresholding or active contour models, giving binary masks that remove non-informative pixels. Hair removal is then added to reduce occlusions that may alter texture and spectral characteristics. Morphological filtering and DullRazor-like inpainting are used to fill the skin-colored pixels where dark hair artifacts are present. Lastly, hue-gamma correction balances the color and luminance contrast to increase the visual separability of benign and malignant patterns of lesions. This step is of special importance in highlighting slight differences in pigmentation and structural anomalies.

Background Removal: Dermoscopic images may include non-skin background areas that may pose bias to feature extraction. To remove this, a binary segmentation mask is formed through thresholding or clustering of the grayscale image as given in eqn. (8),

$$Mx,y=1\text{if } I(x,y)\geq T0\text{otherwise} \quad (8)$$

where I(x,y) represents intensity of the image at the pixel (x,y) and the adaptive threshold is T. The area of the lesion is then obtained as eqn. (9),

$$Ilesionx,y=Ix,y.M(x,y) \quad (9)$$

This isolates the lesion, so that the background does not influence feature learning.

Hair Removal with a Morphological Filtering: The boundaries of lesions are often obscured by hair structures and this results in misclassification. Morphological closing and inpainting are used. To begin with, hair artifacts are detected by a top-hat transform as given in eqn. (10),

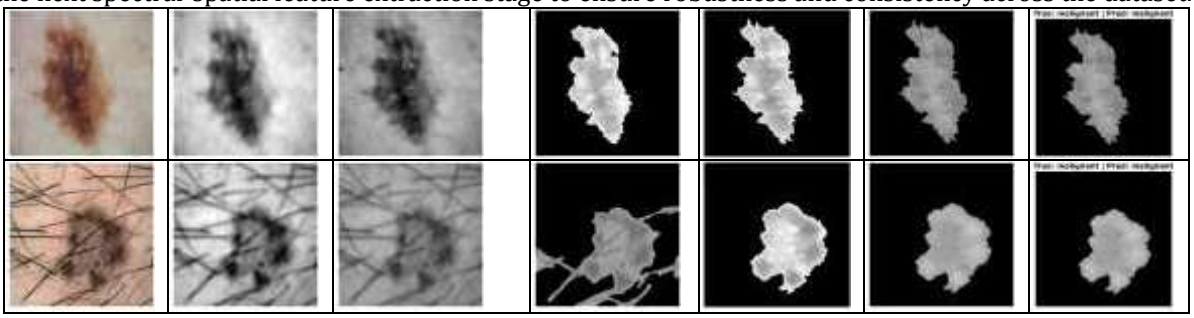
$$Hx,y=Ix,y-(I\circ S)(x,y) \quad (10)$$

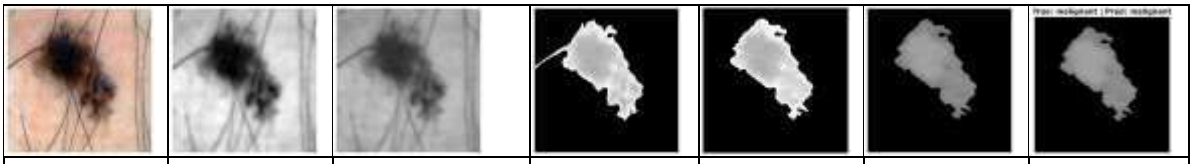
where S is a structuring element, the morphological opening is denoted by ∘. The lesion region is then reconstructed without artifacts by inpainting H(x,y) with neighbouring pixel values.

Contrast Enhancement by Hue-Gamma Correction: The uneven light and low color contrast can hide the features of melanoma. The preservation of hue, coupled with gamma correction, makes lesions more visible. Gamma correction is specified using eqn. (11),

$$Ioutx,y=Iin(x,y)\gamma \quad (11)$$

where γ<1 defines the darker areas are brightened (to show the structures behind them). γ>1 representing the overly bright areas are suppressed by the use of the symbol gamma greater than one. In dermoscopic images, values are usually between 0.8≤γ≤1.2. This procedure enhances the contrast between the malignant and benign lesion areas, which makes feature extraction more reliable. The resulting pre-processed images, which are free of confounding artifacts and enriched in lesion contrast, serve as the standardized inputs to the next spectral-spatial feature extraction stage to ensure robustness and consistency across the dataset.

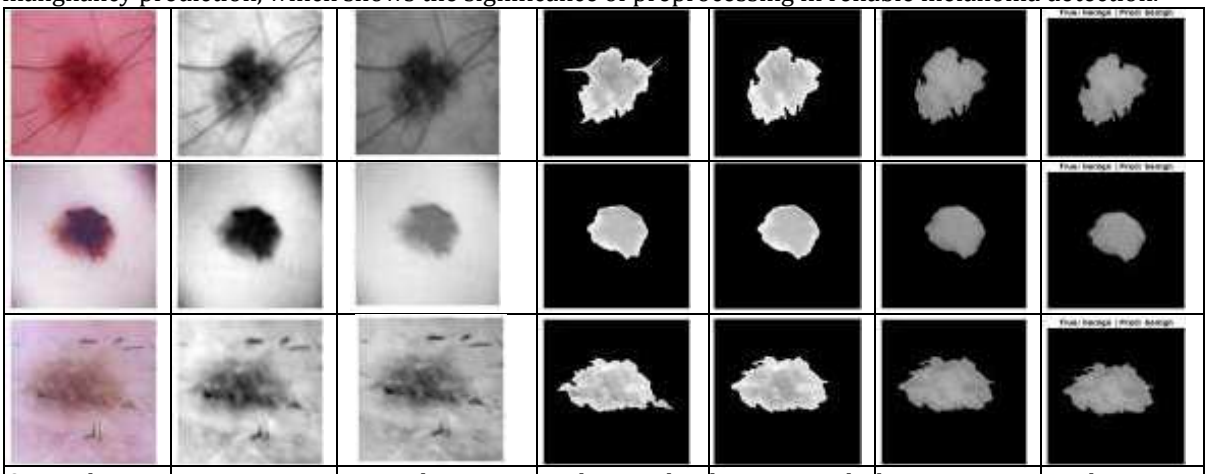




Original	IRT	spectral decomposition	Background removal	hair removal	hue-gamma correction	Prediction

Figure 2: Preprocessing pipeline for malignant melanoma detection

The figure 2 illustrates the preprocessing procedure used on malignant melanoma images prior to classification. The original dermoscopic image is subjected to Infrared Thermography (IRT) transformation and spectral decomposition to bring out the concealed structures of the lesion. Background removal isolates the area of the lesion, removing irrelevant areas. Morphological filtering to remove hair enhances the appearance of lesion boundaries. Hue-gamma correction increases contrast, which provides a clearer visualization of irregular pigmentation. These measures enhance the quality of images and highlight the melanoma-specific patterns. The processed image is then fed in the deep learning model leading to accurate malignancy prediction, which shows the significance of preprocessing in reliable melanoma detection.



Original	IRT	spectral decomposition	Background removal	hair removal	hue-gamma correction	Prediction

Figure 3: Preprocessing pipeline for Benign melanoma detection

The figure 3 shows the preprocessing pipeline of benign melanoma detection, which indicates the transformation stages that are necessary to enhance image clarity and classification accuracy. The original dermoscopic image is then transformed using Infrared Thermography (IRT) and spectral decomposition, which makes the lesion more visible and show minor differences in structure. Background removal subsequently isolates the lesion against the surrounding skin, so that non-relevant areas do not affect the feature extraction. Morphological filtering effectively removes hair artifacts that are usually common in dermoscopic scans, and this does not distort the boundaries of the lesion. Hue-gamma correction also enhances contrast and color homogeneity, highlighting the benign lesion features of smooth borders and homogeneous pigmentation. These preprocessing steps make sure that the region of lesion is well represented before it is fed into the deep learning classifier. The resulting prediction is a clear indication of the lesion being benign, showing how a preprocessing pipeline can be used to increase robustness and reduce noise as well as allow reliable classification between benign and malignant melanoma in an automated detection system.

3.5 Feature extraction

Spectral-spatial features are extracted from the pre-processed hyperspectral-like images. These features capture both spectral intensity variations and spatial lesion morphology.

Spectral Feature Representation: For each pixel (x,y), the spectral vector is mathematically expressed using eqn. (12),

$$Sx,y=I\lambda 1x,y,I\lambda 2x,y,\ldots,I\lambda bx,y \quad (12)$$

where $I\lambda ix,y$ is the intensity at band λi and b is the number of bands.

Spatial Texture Features: Local spatial descriptors are obtained by applying convolution with spatial filters fk and it is mathematically expressed using eqn. (13),

$$Tkx,y=u=-mmv=-mmIx+u.y+vfk(u,v) \quad (13)$$

where fk represents a texture kernel (e.g., Gabor, LoG), and mmm defines the neighbourhood size.

Spectral–Spatial Fusion Feature Vector: The joint spectral–spatial feature vector is then formed using eqn. (14),

$$Fx,y=\alpha.Sx,y\oplus\beta.T(x,y) \quad (14)$$

where $\oplus$ denotes concatenation, and α, β are weighting parameters balancing spectral vs. spatial contributions. Moreover, the extracted features were fused before selecting the optimum features.

3.6 Feature Selection

The feature selection is carried out based on the Whale Optimization Algorithm (WOA) in eqn. (14). WOA is a metaheuristic that mimics the social hunting behaviour of humpback whales, particularly their bubble-net feeding technique. It effectively eliminates redundancy by picking the most relevant subset of features out of the high-dimensional fused vector. WOA is based on the bubble-net hunting mechanism of humpback whales, in which candidate feature subsets are iteratively updated using exploitation (shrinking encircling prey, spiral updating) and exploration (random search of global optima). The fitness was defined as the classification accuracy and feature subset size to get a trade-off between performance and dimensionality reduction. Lastly, the best feature set identified by WOA was the one that kept the most discriminative features, which enhanced the accuracy of melanoma detection and minimized the computational complexity.

Phase 1: Initialization

A population of whales (candidate feature subsets) is randomly initialized from the fused feature set Fx,y. Each whale position represents a potential feature subset. The dimensionality of each whale corresponds to the length of the fused vector. Moreover, feature initialization has been given in eqn. (15),

$$Xi=\{f1,f2,\ldots,fd\},\ i=1,2,\ldots,N \quad (15)$$

where Xi denotes the feature subset of the i-th h whale, d is the dimension of the feature vector, and N is the population size. The fitness function evaluates each subset based on classification accuracy (Acc) and the number of selected features (|X|), given in eqn. (16),

$$FitnessX=\lambda.AccX-(1-\lambda)|X|d \quad (16)$$

where $0\leq\lambda\leq1$ controls the trade-off between accuracy and subset size.

Phase 2: Encircling Prey (Exploitation)

Whales presume that the best solution that has been identified is the prey (optimal subset). The encircling mechanism revises the positions around the best solution X* and it has been mathematically expressed using eqn. (17)-(18),

$$D=|C.X*t-Xt| \quad (17)$$

$$Xt+1=X*t-A.D \quad (18)$$

Where, $A=2a\cdot r-a$, $C=2r$, $r\in[0,1]$ representing the random vector is and a decreases linearly from 2 to 0 over iterations. This allows whales to shrink their search around the best solution.

Phase 3: Bubble-Net Attacking (Exploitation - Spiral Updating)

Whales also rotate positions in a spiral-like way to simulate bubble-net hunting and it is mathematically expressed using eqn. (19),

$$Xt+1=D'.ebl.cos2\pi l+X*t \quad (19)$$

where $D'=X*(t)-X(t)$, b is a constant defining the logarithmic spiral's shape, and $l\in[-1,1]$ is a random number. This balances local exploitation by allowing whales to converge non-linearly around the prey.

Phase 4: Exploration (Global Search)

If $|A|\geq1$, whales randomly move relative to another whale's position, enabling exploration and avoiding premature convergence and it is mathematically expressed using eqn. (19)(20),

$$D=C.Xrand-X(t) \quad (19)$$

$$Xt+1=Xrand-A.D \quad (20)$$

where Xrand is a randomly selected whale from the population.

Phase 5: Optimal Feature Subset Selection

The algorithm iterates through these phases until a stopping criterion (maximum iterations or convergence) is met. The best whale position X* represents the optimal subset of spectral–spatial features is mathematically expressed using eqn. (21),

$$Foptx,y\subseteq F(x,y) \quad (21)$$

This subset maximizes discriminative power for melanoma detection while minimizing redundancy and computational cost.

The Whale Optimization Algorithm (WOA) feature selection workflow, shown in Figure 4, shows the methodical procedure of obtaining an optimal set of features out of the high-dimensional spectral-spatial fusion feature vector. This is initiated by a random generation of a population of whales (candidate feature subsets). Each whale is a binary string that denotes whether a feature is selected or not the algorithm next

enters the encircling prey stage (exploitation), whereby whales move towards the current best solution by changing their positions relative to the leader. This guarantees convergence to promising feature subsets

After that, the bubble-net attacking mechanism presents a spiral updating model, which is a simulation of the bubble-net feeding strategy of the whales. This mechanism improves the search locally and enables the best solution to be exploited precisely. In the meantime, the exploration process (global search) avoids early convergence by stimulating whales to move randomly among the other candidates, increasing diversity and avoiding local minima. During these iterative updates, a fitness criterion based on the accuracy of the classification and the number of selected features guides the selection process. Lastly, the algorithm determines the best feature subset that can balance accuracy and dimensionality reduction. The workflow renders WOA very applicable to spectral-spatial fusion data, where the feature selection is compact but discriminative in order to achieve better performance in classification.

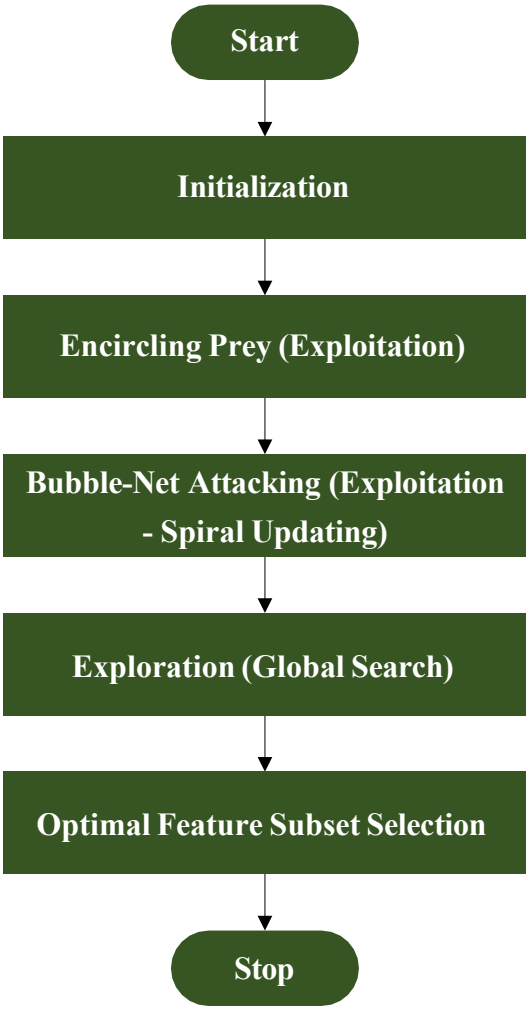


Figure 4: Workflow of WOA

Algorithm 1: Pseudocode for Feature selection using WOA

Input:
F(x,y) = Spectral–Spatial Fusion Feature Vector (Eqn. 14)
N = Number of whales (population size)
Tmax = Maximum iterations
d = Dimension of feature space
ω = Weighting factor (accuracy vs. feature count)
Output:
F* = Optimal feature subset
Begin
Initialize whale population Xi=(1,2,…,N) with random binary feature subsets

Initialize parameters a = 2, A, C
For each Xi do
Evaluate fitness using:
FitnessX=λ.AccX-(1-λ)|X|d
End For
Identify best solution X* with minimum fitness
For t = 1 to Tmax do
Update parameter:
a = 2 - (2 * t / Tmax)
For each whale Xi do
Generate random numbers p, r in [0,1]
Compute A=2a·r-a
Compute C=2r
If p < 0.5 then
If |A| < 1 then
Encircling prey (exploitation)
D=|C.X*t-Xt||
Xt+1=X*t-A.D
Else
Exploration (search for prey)
Select random whale Xrand
D=C.Xrand-X(t)
Xt+1=Xrand-A.D
End If
Else
// Spiral updating (bubble-net behavior)
D' = |X* - Xi|
l = random number in [-1,1]
Xi=D'.ebl.cos2πl+X*t
End If
Apply sigmoid transfer function + threshold to convert Xi to binary {0,1}
Evaluate fitness of updated Xi
End For
Update best solution X* if better fitness is found
End For
Return X* as the optimal feature subset F*
End

3.7 Classification

Then, the selected features are fed into NASceptionNet, a hybrid deep learning model combining NASNet and XceptionNet. This model enables robust lesion classification into Malignant or Benign categories. After obtaining the best feature set Fopt(x,y), then Whale Optimization Algorithm (WOA) is used to obtain the optimal spectral-spatial features (x,y), which are then fed into the proposed NASceptionNet classifier. This combined deep learning model combines NASNet (Neural Architecture Search to automatically extract features) and XceptionNet (depthwise separable convolutional efficiency). The aim is to take advantage of the automated design feature of NASNet and the lightweight and powerful spatial encoding of XceptionNet. Mathematically, the input mapping to the hybrid network is expressed using eqn. (22),

$$Z=\phi(Foptx,y) \quad (22)$$

Here, Foptx,y is the optimal spectral–spatial feature set, ϕ(.) represents preprocessing operations such as normalization and reshaping, and Z is the structured input tensor for NASceptionNet. The NASNet feature extractor applies learned convolutional cells, producing higher-order representations and it is mathematically represented as eqn. (23),

HNAS=CNAS(Z) (23)
Similarly, the XceptionNet module applies depthwise separable convolutions CXcep, which efficiently capture fine-grained lesion structures and it is mathematically expressed using eqn. (24),
HXcep=CXcep(HNAS) (24)
The hybrid representation is then fused by concatenation or weighted addition and it is mathematically expressed using eqn. (25),
Hfusion=λ1HNAS⊕λ2HXcep (25)
where λ1, λ2 are learnable fusion weights, and ⊕ indicates feature fusion. Finally, lesion classification into Malignant (M) or Benign (B) categories is performed using a fully connected layer with a softmax activation as given in eqn. (26),
PcFopt=Softmax(W.Hfusion+b) (26)
Here, W,b are trainable parameters, PcFopt gives the class probability distribution over {M,B}. Thus, the predicted lesion type is obtained using eqn. (27),
c=argmaxc∈{M,B}PcFopt (27)
Final Decision Rule is based on **t**he predicted class label c is determined by maximum probability as given in eqn. (28),
c=Malignantif PMFopt>PBFoptBenignOtherwise (28)

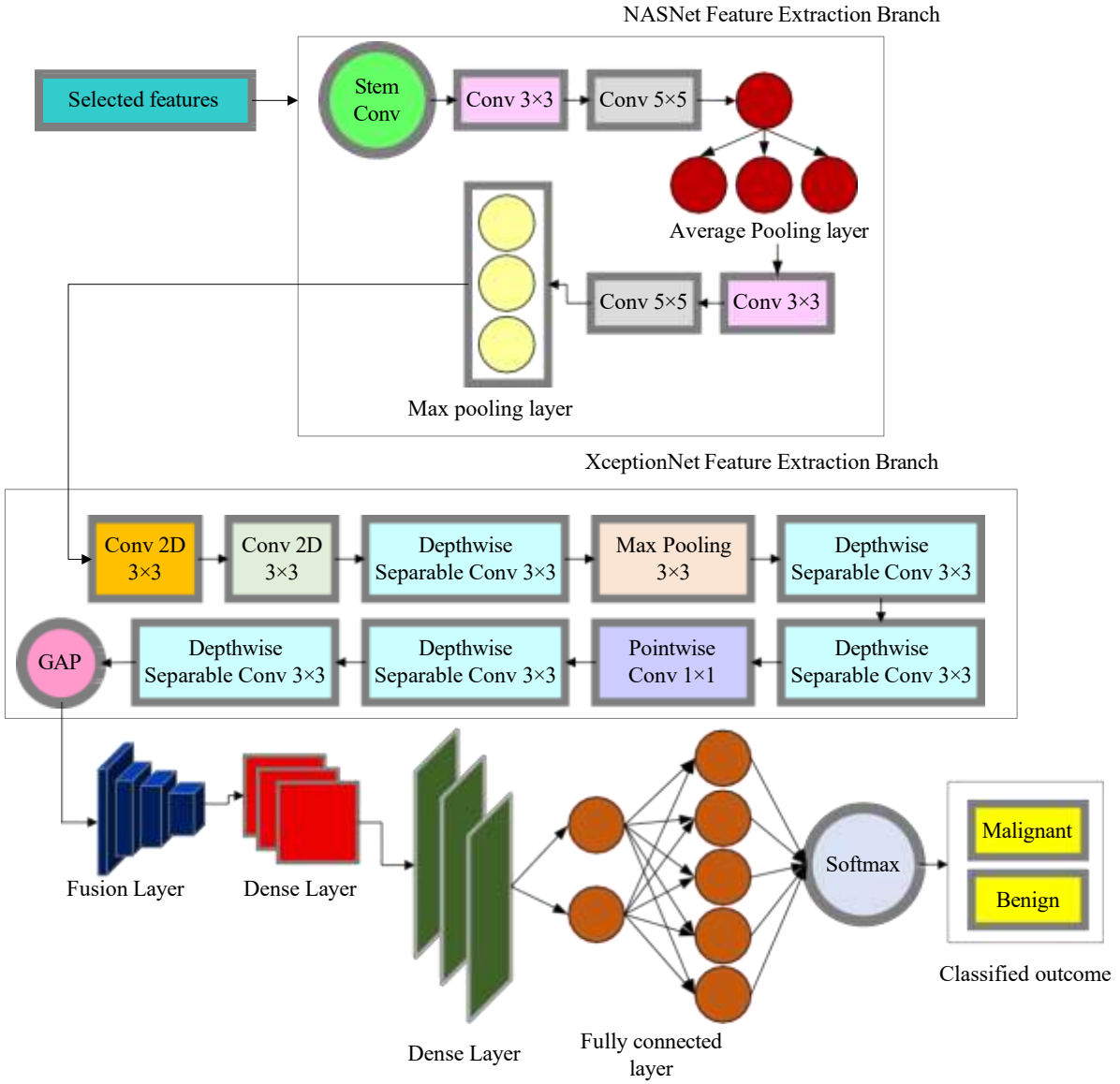


Figure 5: Architecture for NASceptionNet

The proposed NASceptionNet hybrid classification model in Figure 5 integrates the advantages of NASNet and XceptionNet to obtain strong lesion classification as Malignant or Benign. The pipeline starts with the chosen features of spectral-spatial fusion and WOA-based feature selection. The NASNet branch processes

these features with a stem convolution, a series of normal cells that include depthwise separable convolutions, pooling layers, and skip connections. This architecture can extract multi-scale feature representation and minimize redundancy. The downsampling cell makes sure that downsampling is effective, creating small feature maps. The XceptionNet branch simultaneously learns high-level discriminative features. The entry flow is composed of standard convolution and max pooling layers, and the middle flow is composed of depthwise separable convolutions with pointwise convolution and skip connections, which enhance the efficiency and avoid overfitting. The exit flow uses additional separable convolutions and global average pooling (GAP) to reduce feature maps to a representative vector. The fusion layer merges the outputs of NASNet and XceptionNet to produce a hybrid representation. This is then followed by dense layers that fine-tune discriminative patterns and also reduce dimensionality. The last layer is fully connected with a softmax activation that classifies lesions as malignant or benign. This hybrid design is more accurate, efficient, and robust because of the complementary feature extraction strategies employed.

4. Result and Discussion

This section provides an overall performance carried out by the proposed model. For validating the performance of the suggested model, the performance score of the proposed model has been measured and compared with the existing models. Furthermore, the performance metrics such as Accuracy, Precision, Recall, F1-Score, Specificity, NPV, MCC, FPR and FNR respectively. Apart from these this section also delves the dataset description along with experimental setup of the study.

4.1 Dataset description

The Skin Cancer: Malignant vs. Benign dataset, curated by Claudio Fanconic and sourced from the ISIC Archive, consists of dermatological images categorized by lesion type (benign or malignant) available in both training and testing sets. It includes 1,800 benign and 1,497 malignant images — a nearly balanced collection. All images are standardized to 224×224 pixels in RGB format, making them readily usable for image classification tasks such as training Convolutional Neural Networks (CNNs) Kaggle+1. The labelled separation and uniform formatting make this dataset particularly suitable for developing and evaluating deep-learning models aimed at differentiating between benign and malignant skin lesions.

4.2 Experimental setup

The experimental system combines preprocessing, spectral simulation, optimization, and hybrid deep learning to detect melanoma robustly. The Kaggle dataset is a balanced collection of malignant and benign images that are pre-processed and converted into IRT-simulated thermal maps. These maps are broken down into pseudo-hyperspectral bands, which contain more information about the lesions than RGB. Whale Optimization Algorithm guarantees that only the most significant spectral-spatial features are kept to decrease redundancy. NASNet, which is a combination of NASNet and Xception blocks, has a good representational power and efficient classification. The effectiveness of the model is confirmed by multiple metrics, such as F1 and AUC, which guarantees the effectiveness of the framework to screen melanoma at an early stage. Moreover, the following table 2 delves the Experimental Setup for IRT-Simulated Hyperspectral Melanoma Detection.

Table 2: Experimental Setup for IRT-Simulated Hyperspectral Melanoma Detection

Component	Description
Dataset	Kaggle Skin Cancer: Malignant vs. Benign (3,297 images)
Input Processing	Resize to 224×224, background & hair removal, gamma correction
IRT Simulation	Synthetic steady-state & rewarming thermal maps
Hyperspectral Gen.	12 pseudo-spectral bands via spectral decomposition
Feature Selection	Whale Optimization Algorithm (30 population, 50 iterations)
Classifier	NASceptionNet (NASNet + Xception hybrid)
Training	AdamW, lr=1e-4, batch=32, 60 epochs

4.3 Metrics analysis

Table 3 shows the key indicators of the melanoma detection performance. Accuracy gives the general correctness, but it is misleading in one-sided datasets. Precision and NPV evaluate the trustworthiness of positive and negative predictions, respectively, whereas Recall and Specificity reflect the sensitivity of the model to the true cases and the avoidance of false alarms. The F1-Score is a good compromise between Precision and Recall, and is therefore useful in clinical settings. MCC provides a global evaluation, taking into consideration all the elements of the confusion matrix. FPR and FNR reflect important error rates and can be used to gain insight into false diagnoses. In combination, these metrics provide balanced and trustworthy model assessment.

Table 3: Analysis of performance metrics

Metric	Description	Formula
Accuracy	Proportion of correctly classified cases (overall correctness).	A=tp+tntp+fp+tn+fn×100
Precision	Proportion of predicted positives that are truly positive (reliability of positive prediction).	P=tptp+fp×100
Recall (Sensitivity/TPR)	Proportion of actual positives correctly detected (ability to find melanoma).	R=tptp+fn×100
F1-Score	Harmonic mean of Precision and Recall (balance measure).	F1-score=2×precision×Recallprecision+Recall×100
Specificity (TNR)	Proportion of actual negatives correctly identified (ability to avoid false alarms).	SP=tntn+fp×100
NPV (Negative Predictive Value)	Proportion of predicted negatives that are truly negative.	NPV=tntn+fn×100
MCC (Matthews Correlation Coefficient)	Balanced measure considering all outcomes (−1 to +1).	MCC=tp×tn-(fp×fn)(tp+fp)(tp+fn)(tn+fp)(tn+fn)
FPR (False Positive Rate)	Proportion of negatives incorrectly classified as positive.	FPR=fpfp+tn×100
FNR (False Negative Rate)	Proportion of positives incorrectly classified as negative.	FNR=fnfn+tp×100

4.4 Performance comparison analysis

Table 4 shows the comparison of the performance of the proposed melanoma detection framework before and after IRT-based hyperspectral conversion. The findings indicate that the IRT transformation significantly improves the classification results in all the assessment measures. Accuracy was increased by 1.08% to 99.12%, which demonstrates the enhanced reliability of the model in the correct classification of both malignant and benign lesions. Precision and Recall also increased, to 99.05 and 99.18 respectively, which means that the system did not only reduce false positives, but also managed to capture more true cases of melanoma. In medical diagnostics, this is especially important to reduce false negatives (FNR), as it is important not to miss malignant lesions.

Table 4: Performance of the proposed model as before vs after IRT conversion

Metric	Before IRT conversion	After IRT conversion
Accuracy	98.04	99.12
Precision	97.88	99.05
Recall	98.21	99.18
F1-Score	98.04	99.11
Specificity	97.87	99.06
NPV	98.21	99.18
MCC	96.09	98.23
FPR	2.13	0.94
FNR	1.79	0.82

The F1-Score increased to 99.11% as compared to 98.04%, indicating a more balanced sensitivity and reliability. Specificity improved to 99.06%, so benign lesions were better classified, and fewer false alarms were generated. In like manner, the Negative Predictive Value (NPV) increased, which further supports the reliability of the model in excluding melanoma. The Matthews Correlation Coefficient (MCC), which takes into consideration all elements of the confusion matrix, increased significantly, by 2.14%, to 98.23%,

indicating better overall predictive performance. The error rates also demonstrate the importance of the IRT conversion: the False Positive Rate (FPR) decreased, being 2.13% and 0.94%, respectively, and the False Negative Rate (FNR) was 1.79% and 0.82%, respectively. These decreases are clinically important because fewer benign cases are missed as malignant and fewer true melanomas are missed. In short, the IRT conversion can effectively boost spectral-spatial representation, which allows more effective feature extraction and increases the accuracy, reliability, and clinical applicability of the model used to detect melanoma in its early stages. Following Figure 6 shows the Graphical representation of Performance comparison of the proposed model as before vs after IRT conversion.

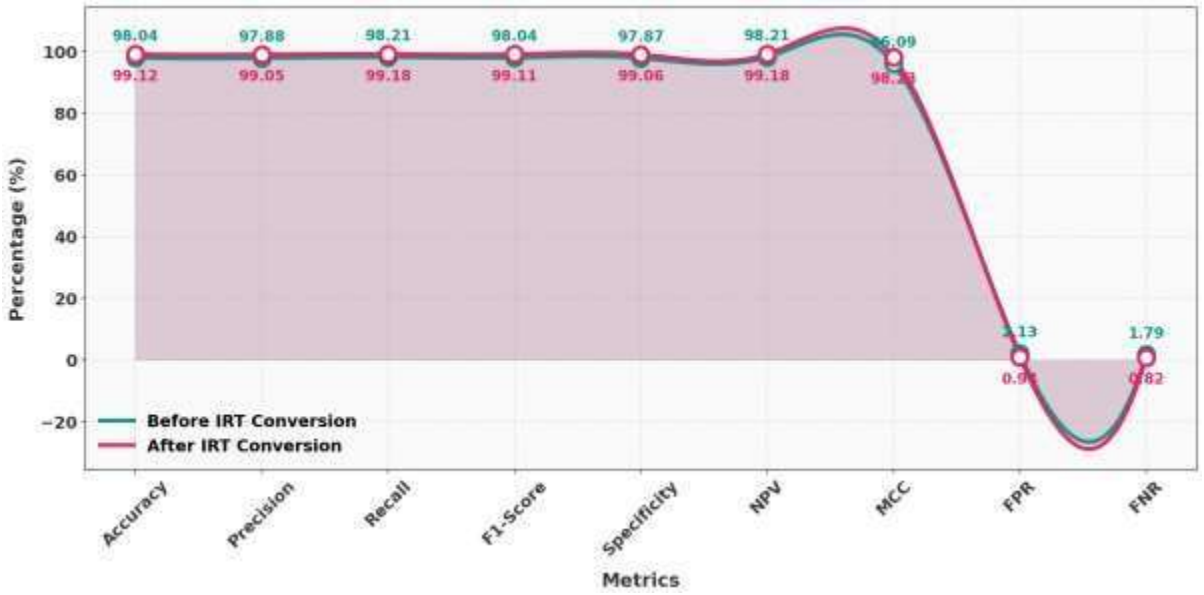


Figure 6: Graphical representation of Performance comparison of the proposed model as before vs after IRT conversion

Table 5 shows the comparison of the proposed IRT-simulated hyperspectral deep learning model with some of the state-of-the-art methods of melanoma detection. The findings are clear in that the proposed framework is superior especially following the IRT conversion. The traditional models included Explainable DL [13], Hybrid CNN [15], and Hybrid CAD [16], which had an accuracy of 93.4-96.3. Although these models offer good baselines, they have relatively lower Recall and MCC scores, which implies that they are not very effective in capturing all the melanoma cases and maintaining balanced predictive performance. The performance of MiSC [19] and DSC-EDLMGWO [20] was better, with DSC-EDLMGWO achieving an accuracy of 98.38% and good overall measures. Nevertheless, the suggested model is even better compared to the most powerful rivals. It already had a competitive accuracy of 98.04% with a balanced Precision (97.88%) and Recall (98.21%) prior to IRT conversion. With the introduction of IRT-based spectral transformation, the performance was improved even more in all dimensions. Accuracy was 99.12%, with Precision and Recall above 99% so that false positives and false negatives were minimal. The F1-Score increased to 99.11% which further proved that the sensitivity and reliability were balanced. Notably, the MCC was improved to 98.23, which is significantly higher than the 96.9 obtained by DSC-EDLMGWO, demonstrating a greater robustness of all the confusion matrix results.

Table 5: Comparison of proposed model along with state-of-art models

Model / Method	Accuracy (%)	Precision	Recall	F1-Score	Specificity	NPV	MCC	FPR	FNR
Explainable DL [13]	95.23	95	94.8	94.9	95.6	95.1	90.7	4.4	5.2
Hybrid CNN [15]	96.3	95.7	94.4	95	96.5	96.2	91.8	3.5	5.6
Hybrid CAD [16]	93.4	93	92.5	92.7	93.8	93.1	87.2	6.2	7.5
MiSC [19]	95.6	96.8	95.6	95.7	95.9	95.8	91.5	4.1	4.4

DSC-EDLMGWO [20]	98.38	98.4	98.1	98.2	98.6	98.3	96.9	1.4	1.9
Before IRT Conversion	98.04	97.88	98.21	98.04	97.87	98.21	96.09	2.13	1.79
After IRT Conversion	99.12	99.05	99.18	99.11	99.06	99.18	98.23	0.94	0.82

The advantages of IRT conversion are also evidenced by error rates. The False Positive Rate decreased to 0.94%, whereas the False Negative Rate decreased to 0.82%. These enhancements are clinically important, since they will minimize false alarms on benign lesions and minimize the number of melanomas missed. The proposed model shows not only better statistical results but also better clinical applicability, as the minimization of misclassifications is the key issue in early cancer detection. In general, the combination of IRT-simulated hyperspectral conversion with NASceptionNet and WOA-based feature optimization is a strong improvement, outperforming the current state-of-the-art models in terms of diagnostic accuracy and reliability. Graphical representation of state-of-art comparison is shown in Figure 7.

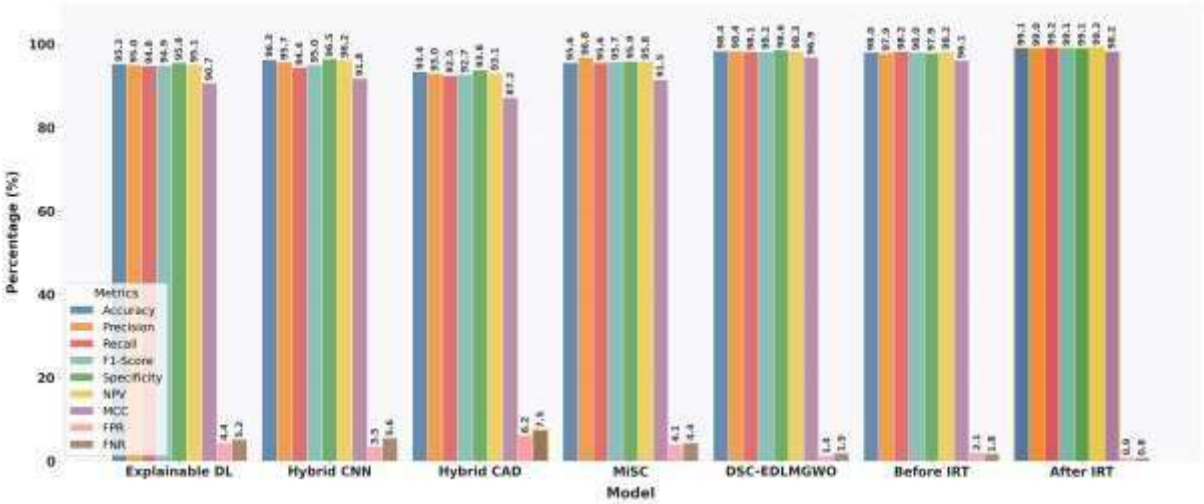


Figure 7: Graphical representation of state-of-art comparison

The use of CNN, ANN, RNN, LSTM, and DNN as the baseline models will give an in-depth comparison of conventional deep learning methods in melanoma detection. Convolutional Neural Networks (CNNs) [21] are popular in medical imaging due to their capacity to extract spatial patterns and local features using convolutional filters, which makes them an obvious choice in dermoscopic image analysis. Artificial Neural Networks (ANNs) [22] are simpler but they provide a benchmark because they model non-linear relationships between input features and outputs and therefore provide a baseline to performance. Recurrent Neural Networks (RNNs) [23] are also added due to their capability to learn sequential dependencies that can be modified to learn pixel or spectral sequences in medical imaging. LSTMs are an improved version of RNNs that are not susceptible to the vanishing gradient problem and are thus more effective at learning long-term dependencies, which makes them appropriate when analyzing spectral or temporal variations in lesion images. Lastly, Deep Neural Networks (DNNs) [24] offer a deeper, fully connected network that can learn complex feature representations and can be seen as a powerful conventional alternative to hybrid and optimized models. Collectively, these models form a well-balanced continuum of conventional deep learning models, which can serve as useful benchmarks against which the proposed IRT-simulated hyperspectral model can be critically compared.

Table 6 shows a detailed comparison of the proposed IRT-simulated hyperspectral deep learning framework and some of the conventional deep learning models such as CNN, ANN, RNN, LSTM, and DNN. The discussion points out the benefits of combining IRT conversion and advanced feature optimization with NASceptionNet to detect melanoma. The accuracies of conventional models CNN and ANN were 95.4 and 94.6 respectively. Although these models show good baseline performance, they are not very sensitive and robust. CNN had a high precision (95.2%) and a slightly lower recall (94.8%), indicating that a few cases of malignancy may be overlooked. NN was a little bit worse in all metrics, having an MCC of 92.7, which is

indicative of the inability to balance the true positives and true negatives. These findings indicate the danger of using only traditional architecture without sophisticated spectral-spatial representation. The results were better in recurrent based models, i.e., RNN and LSTM. RNN had an accuracy of 95.8 % and balanced results on the metrics, although its False Positive Rate was relatively high (4.9 %). LSTM outperformed, with an accuracy of 96.5, precision of 96.2 and recall of 96.4. The lower FNR (3.6%) renders LSTM more dependable in detecting melanoma cases than CNN and ANN. Likewise, the DNN had the best results of the traditional models, with 96.9 accuracy and an MCC of 95.9. Nonetheless, even LSTM and DNN were not robust enough to be used in clinical-level applications, where even a single misclassification may result in life-threatening outcomes.

Table 6: Comparison of proposed model with the traditional DL models

Model	Accuracy (%)	Precision	Recall	F1-Score	Specificity	NPV	MCC	FPR	FNR
CNN	95.4	95.2	94.8	95	94.9	95.1	93.8	5.1	5.2
ANN	94.6	94.3	94.1	94.2	94	94.2	92.7	6	5.9
RNN	95.8	95.6	95.2	95.4	95	95.5	94.1	4.9	4.8
LSTM	96.5	96.2	96.4	96.3	96.1	96.2	95.5	3.9	3.6
DNN	96.9	96.5	96.7	96.6	96.4	96.6	95.9	3.6	3.3
Before IRT Conversion	98.04	97.88	98.21	98.04	97.87	98.21	96.09	2.13	1.79
After IRT Conversion	99.12	99.05	99.18	99.11	99.06	99.18	98.23	0.94	0.82

The proposed model, without IRT conversion, had an accuracy of 98.04% and MCC of 96.09 which is better than traditional models. This illustrates that the hybrid NASceptionNet and Whale Optimization Algorithm (WOA)-based feature selection provide a better baseline. The precision (97.88) and recall (98.21) values show that the model performed better in terms of balancing positive identification reliability and sensitivity than conventional DL models. It is worth noting that the FNR decreased to 1.79% which is much lower than that of CNN (5.2%), ANN (5.9%), and even DNN (3.3%). This decrease is critical as undetected cases of melanoma can have a direct impact on patient outcomes. The performance of the proposed model after IRT conversion was even better and this demonstrated a clear advantage of the proposed model over the traditional methods. The accuracy increased to 99.12%, which is more than 2% higher than the best-performing DNN. Precision and recall were above 99%, and it was confirmed that both malignant and benign lesions were classified with high reliability. The F1-score was 99.11%, which is a high indicator of the sensitivity-precision balance. Specificity and NPV also increased, decreasing false alarms and increasing confidence in negative predictions. Notably, the MCC increased to 98.23, which shows that the model was able to perform well in all the classification results even in the difficult cases. The clinical significance of the IRT conversion is illustrated by the error rates. The FPR was lowered to 0.94 percent and FNR to 0.82 percent, which is far much lower than all the conventional models. This is a breakthrough in the accuracy of the diagnosis compared to CNN (FPR 5.1%, FNR 5.2%) and ANN (FPR 6.0%, FNR 5.9%). Thus, the following Figure 8 shows the Graphical representation of proposed model comparison along with the traditional DL models.

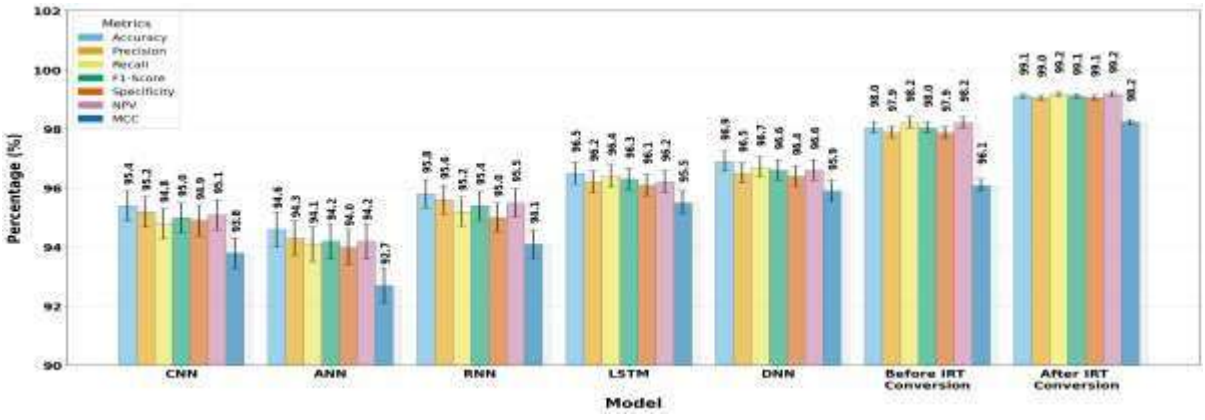


Figure 8: Graphical representation of proposed model comparison along with the traditional DL models

The proposed framework shows a better reliability than even advanced LSTM and DNN models in avoiding misdiagnosed benign cases as well as missed melanoma cases. In sum, Table 6 shows that conventional deep learning models are helpful but have limitations in dealing with the complexity of melanoma detection. The suggested IRT-simulated hyperspectral deep learning framework can greatly improve performance through enrichment of spectral-spatial information, optimization of features, and use of a powerful hybrid classifier. Its high and stable performance and low error rates make it a potential clinically viable, non-invasive diagnostic tool in early detection of melanoma, establishing a new standard compared to the conventional deep learning models.

4.5 Discussion

The proposed study presents an IRT-simulated hyperspectral deep learning framework to detect melanoma accurately, although there are some limitations that need to be considered. To begin with, the data employed (Kaggle Skin Cancer: Malignant vs. Benign) has quite small sample sizes compared to those in the real-life clinical practice, which can influence the applicability of the model to various populations and imaging situations. The simulated IRT conversion is good but it might not be exactly like real thermal imaging and experimental validation with actual IRT data would make the results stronger. Also, the computational demands of spectral decomposition and Whale Optimization Algorithm (WOA)-based feature selection are not low, which may restrict the deployment of the method in low-resource clinical settings. Nevertheless, the study is of great contribution. It shows that combining IRT simulation with hyperspectral decomposition can give more spectral-spatial information, which can be used to improve the accuracy of classification compared to conventional RGB-based methods. The proposed hybrid NASceptionNet with WOA optimization exhibits a strong performance and better error reduction than the state-of-the-art and conventional deep learning models. In clinical terms, the decrease in false negative and false positive is extremely important, as it reduces the number of missed cases of melanoma and unnecessary interventions. On the whole, the research is a positive contribution to the creation of non-invasive, reliable, and efficient melanoma screening tools that would fill the gap between the sophisticated spectral analysis and the practical medical application.

5. Conclusion

This paper presented an IRT-simulated hyperspectral deep learning model to detect melanoma, which combined spectral-spatial feature extraction, Whale Optimization Algorithm (WOA)-based feature selection, and a hybrid NASceptionNet classifier. The findings clearly show the effectiveness of the proposed model over the traditional deep learning and state-of-the-art methods. IRT conversion resulted in an accuracy of 99.12 percent, which was better than the pre-IRT version (98.04 percent) by 1.1 percent. Precision increased by 1.2 percent (97.88 to 99.05), Recall by 1 percent (98.21 to 99.18), and F1-score by 1.1 percent (98.04 to 99.11). Specificity and NPV improved by about 1.2%, whereas MCC improved by 2.2% (96.09 to 98.23). Notably, the error rates were drastically decreased, with FPR decreasing by 55.8% (2.13% to 0.94%) and FNR by 54.1% (1.79% to 0.82%). These enhancements confirm the clinical relevance of the model, and it can be used as a non-invasive method of early melanoma detection and a promising approach to the use of spectral deep learning in medical diagnostics.

5.1 Future scope

The proposed IRT-simulated hyperspectral deep learning model shows great potential in the detection of melanoma; nevertheless, there are several areas that could be improved to make the model more applicable. Second, the model could be enhanced by increasing the size of the dataset, using multi-institutional and ethnically diverse samples, which will enhance the generalizability of the model to real-life clinical settings. The use of real Infrared Thermography (IRT) imaging, rather than simulated data, would be a stronger validation of the spectral-spatial approach and allow multi-modal fusion of dermoscopic, hyperspectral, and thermal modalities. Moreover, the framework can be tuned to real-time applications using lightweight architectures, model pruning and edge-computing integration, and can be used in a handheld diagnostic device or mobile health application. Future work will also consider explainable AI (XAI) methods to present clinician-friendly interpretations of spectral features to enhance trust and adoption in medical practice. The incorporation of tele dermatology systems would enable screening of melanoma remotely, especially in underserved areas. In addition to melanoma, the proposed methodology can be modified to detect other skin diseases or cancers whose spectral and thermal characteristics are markedly different to healthy tissue. Lastly, integrating this methodology with federated learning would allow privacy-preserving training across a variety of hospitals, providing scalability and data security. These future trends highlight the potential of IRT-based spectral deep learning to transform precision healthcare.

Compliance with Ethical Standards

Conflict of interest

The authors declare that they have no conflict of interest.
Human and Animal Rights
This article does not contain any studies with human or animal subjects performed by any of the authors.
Informed Consent
Informed consent does not apply as this was a retrospective review with no identifying patient information.
Funding: Not applicable
Conflicts of interest Statement: Not applicable
Consent to participate: Not applicable
Consent for publication: Not applicable
Availability of data and material:
Data sharing is not applicable to this article as no new data were created or analyzed in this study.
Code availability: Not applicable

REFERENCES

1. Adegun, A. A., & Viriri, S. (2019). Deep learning-based system for automatic melanoma detection. IEEE Access, 8, 7160-7172.
2. Majtner, T., Yildirim-Yayilgan, S., & Hardeberg, J. Y. (2019). Optimised deep learning features for improved melanoma detection. Multimedia Tools and Applications, 78(9), 11883-11903.
3. Winkler, J. K., Sies, K., Fink, C., Toberer, F., Enk, A., Deinlein, T., ... & Haenssle, H. A. (2020). Melanoma recognition by a deep learning convolutional neural network—performance in different melanoma subtypes and localisations. European Journal of Cancer, 127, 21-29.
4. Daghrir, J., Tlig, L., Bouchouicha, M., & Sayadi, M. (2020, September). Melanoma skin cancer detection using deep learning and classical machine learning techniques: A hybrid approach. In 2020 5th international conference on advanced technologies for signal and image processing (ATSIP) (pp. 1-5). IEEE.
5. Gulati, S., & Bhogal, R. K. (2019, April). Detection of malignant melanoma using deep learning. In International conference on advances in computing and data sciences (pp. 312-325). Singapore: Springer Singapore.
6. Phillips, M., Marsden, H., Jaffe, W., Matin, R. N., Wali, G. N., Greenhalgh, J., ... & Palamaras, I. (2019). Assessment of accuracy of an artificial intelligence algorithm to detect melanoma in images of skin lesions. JAMA network open, 2(10), e1913436.
7. Bisla, D., Choromanska, A., Berman, R. S., Stein, J. A., & Polsky, D. (2019). Towards automated melanoma detection with deep learning: Data purification and augmentation. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops (pp. 0-0).
8. Vinay, B. N., Shah, P. J., Shekar, V., & Vanamala, H. R. (2020, January). Detection of melanoma using deep learning techniques. In 2020 International Conference on Computation, Automation and Knowledge Management (ICCAKM) (pp. 391-394). IEEE.
9. Dascalu, A., & David, E. O. (2019). Skin cancer detection by deep learning and sound analysis algorithms: A prospective clinical study of an elementary dermoscope. EBioMedicine, 43, 107-113.
10. Nahata, H., & Singh, S. P. (2020). Deep learning solutions for skin cancer detection and diagnosis. In Machine learning with health care perspective: machine learning and healthcare (pp. 159-182). Cham: Springer International Publishing.
11. Vijayalakshmi, M. M. (2019). Melanoma skin cancer detection using image processing and machine learning. International Journal of Trend in Scientific Research and Development (IJTSRD), 3(4), 780-784.
12. Hagerty, J. R., Stanley, R. J., Almubarak, H. A., Lama, N., Kasmi, R., Guo, P., ... & Stoecker, W. V. (2019). Deep learning and handcrafted method fusion: higher diagnostic accuracy for melanoma dermoscopy images. IEEE journal of biomedical and health informatics, 23(4), 1385-1391.
13. Mahmud, M. A. A., Afrin, S., Mridha, M. F., Alfarhood, S., Che, D., & Safran, M. (2025). Explainable deep learning approaches for high precision early melanoma detection using dermoscopic images. Scientific Reports, 15(1), 24533.
14. Swapno, S. M. R., Nobel, S. N., Meena, P. K., Meena, V. P., Bahadur, J., & Appaji, A. (2025). Accelerated and precise skin cancer detection through an enhanced machine learning pipeline for improved diagnostic accuracy. Results in Engineering, 25, 104168.
15. Chawla, J., Prabha, C., Balusamy, B., Rakesh, N., & Khan, F. (2025, March). Revolutionizing Skin Cancer Detection: Leveraging Machine Learning and Deep Learning for Precise Lesion Identification. In 2025 International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 1431-1436). IEEE.

16. Kaur, R., GholamHosseini, H., & Linde´n, M. (2025). Advanced deep learning models for melanoma diagnosis in computer-aided skin cancer detection. Sensors, 25(3), 594.
17. Akinrinade, O., & Du, C. (2025). Skin cancer detection using deep machine learning techniques. Intelligence-Based Medicine, 11, 100191.
18. Patil, P. R. (2025). Deep Learning Revolution in Skin Cancer Diagnosis with Hybrid Transformer-CNN Architectures. Vidhyayana-An International Multidisciplinary Peer-Reviewed E-Journal-ISSN 2454-8596, 10(si4).
19. Mohamed, E. H., Abdu, N., Khalil, M., Kamal, H., & Rashed, E. A. (2025). MiSC: A hybrid multi-modal deep learning approach for accurate skin cancer detection: E. Hussein Mohamed et al. Multimedia Tools and Applications, 1-25.
20. Dorathi Jayaseeli, J. D., Briskilal, J., Fancy, C., Vaitheeshwaran, V., Patibandla, R. L., Syed, K., & Swain, A. K. (2025). An intelligent framework for skin cancer detection and classification using fusion of Squeeze-Excitation-DenseNet with Metaheuristic-driven ensemble deep learning models. Scientific Reports, 15(1), 7425.
21. Ashtikar, A., & Shastri, S. (2025). A CNN Model For Skin Cancer Detection And Classification By Using Image Processing Techniques. Journal of Scientific Research and Technology, 251-263.
22. Shah, A., Shah, M., Pandya, A., Sushra, R., Sushra, R., Mehta, M., ... & Patel, K. (2023). A comprehensive study on skin cancer detection using artificial neural network (ANN) and convolutional neural network (CNN). Clinical eHealth, 6, 76-84.
23. Divya, D., & Ganeshbabu, T. R. (2020). Fitness adaptive deer hunting-based region growing and recurrent neural network for melanoma skin cancer detection. International Journal of Imaging Systems and Technology, 30(3), 731-752.
24. Vargas, D. P. M., Ba´ez, H. N., Pe´rez, J. G. B., & Boho´rquez, J. A. C. (2021). Design of a tool for the classification of skin cancer images using Deep Neural Networks (DNN). Ciencia y Tecnologí´a, 65-80.