



RESEARCH ARTICLE

Hybrid ResNet-50 with Gradient-weighted Histogram Technique for Automated Skin Cancer Classification

S. Vadivel Murugan^{1*}, Dr. J. G. R. Sathiaseelan²

Abstract

To develop an accurate and efficient automated system for skin cancer classification based on dermoscopic images and to improve diagnostic quality compared with traditional deep learning classifiers. Dermoscopic images are resized and normalized, then fed into the pre-trained ResNet-50 backbone to extract deep features. The extracted features are then fed into a hybrid classifier with fully connected layers. In addition, a GHT-based feature importance mechanism is implemented to highlight important features and attenuate the influence of less useful ones. The proposed ResGrad Hybrid model achieves a classification accuracy of 89.4%, higher than that of the ResNet-50 model (87%). These results indicate superior performance for skin cancer detection. Here, we provide a hybrid ResGrad approach that combines deep feature extraction with a feature-importance mechanism (derived from GHT) and a feature-selection method, further improving classification performance compared with the standard ResNet-50 approach.

Keywords: Hybrid-ResNet-50, Hybrid Deep Learning Model, Gradient-weighted Histogram Technique (GHT), Feature Importance, Medical Image Analysis, Computer-Aided Diagnosis (CAD), Deep Feature Extraction.

Introduction

Skin cancer is one of the most frequent and serious illnesses that we are experiencing across every continent. Cancerous cells within the skin grow from unregulated growth of abnormal skin tissues, which most often arises from prolonged exposure to ultraviolet (UV) radiation derived from sunlight and artificial tanning devices. Worldwide, skin cancer cases increase in number annually, and early

diagnosis as well as successful treatment are crucial to a reduction in death rates. The medical community identifies two main types of skin cancer: benign tumors that do not spread and malignant tumors that will grow into advanced cancer if not recognized early on.

The conventional diagnosis relies on visual evaluations and meticulous dermatologic examinations by dermatologists, which means that the patient undergoes a lengthy procedure, relying on the subjective judgment of an MD. The health sector requires computer-assisted diagnostic technologies that will yield reliable, timely assessments and allow skin cancer cases to be identified and classified globally, resulting in better outcomes. The skin is the body's largest organ, and it acts as a barrier against infection, heat, and ultraviolet (UV) radiation. Skin cancer does, however, remain one of the most gruesome and very quickly growing health risks globally. Skin cancer is a significant trend, with millions of new cases diagnosed each year in the U.S. Skin cancer happens when there is abnormal change at the levels of the cells in the area of the skin, this makes skin cells grow and divide without the normal control and causes the form the tumors by inappropriate changes of normal growth in other ways of the skin cell structure that cause to the skin to grow and divide abnormally.

The excessive and uncontrolled growth might occur when exposed to too much UV light, smoking, drinking alcohol, environmental changes, viral infections, or chemical

¹Research Scholar, Department of Computer Science, Bishop Heber College (A), Affiliated to Bharathidasan University, Tiruchirappalli-620017, Tamil Nadu, India

²Research Supervisor, Vice Principal and Head, Associate Professor, Department of Computer Science, Bishop Heber College (A), Affiliated to Bharathidasan University, Tiruchirappalli-17, Tamil Nadu, India

***Corresponding Author:** Author, Research Scholar, Department of Computer Science, Bishop Heber College (A), Affiliated to Bharathidasan University, Tiruchirappalli-620017, Tamil Nadu, India, E-Mail: vadivelmurugan462@gmail.com

How to cite this article: Murugan, S.V., Sathiaseelan, J.G.R. (2026). Hybrid ResNet-50 with Gradient-weighted Histogram Technique for Automated Skin Cancer Classification. *The Scientific Temper*, 17(6):6429-6439.

Doi: 10.58414/SCIENTIFICTEMPER.2026.17.6.12

Source of support: Nil

Conflict of interest: None.

abuse. The four primary types of skin cancer can be grouped into two main types, namely benign and malignant. Detection of skin cancer is difficult.

Initially, dermatologists observe suspicious lesions on the skin, and then, after performing imaging analysis via a magnifying microscope (visualizations, patterns), dermoscopy is performed to visualize these lesions to generate a diagnosis.

The diagnosis is usually confirmed through a biopsy. Dermoscopy increases the diagnostic capability, but accurate results still depend on expertise, and manual diagnosis may also be time-consuming and stressful for patients. Computer-aided diagnostic systems have been developed with the aim of solving these challenges, aimed notably to enable these systems to serve clinicians in a field that typically does not feature many specialists. Recent machine learning and deep learning approaches have brought great improvement of the detection and classification of skin cancer.

Recently, convolutional neural networks (CNNs) have attracted great attention for their ability to automate the encoding of meaning from medical images. While there are many studies on benign and malignant lesions to be detected. Early and correct detection remains critical, with earlier treatment having a remarkable impact on the survival and mortality of skin cancer.

Skin cancer is undoubtedly among the most frequent and devastating health problems worldwide, with incidence rates rising every year. It arises from the unregulated proliferation of abnormal skin cells, predominantly arising from long-term exposure to ultraviolet (UV) radiation. Skin tumors can be divided into benign and malignant, accounting for the majority of deaths from skin cancer, although benign, are important in clinical assessment because of their apparent similarity to malignant lesions.

Traditional diagnostic procedures such as visualization, dermoscopy, and biopsy are generally lengthy, dependent on clinical skill, and inevitably yield subjective, heterogeneous outcomes. Dermoscopy increases diagnostic accuracy but does not eliminate uncertainty and thus requires computer-aided diagnostic (CAD) systems. Recently, advances in artificial intelligence (AI), specifically deep learning, have significantly improved classification. Convolutional neural networks (CNNs) such as VGG16, ResNet, DenseNet, and EfficientNet have demonstrated strong feature extraction capabilities for dermoscopic images. Several studies demonstrate that hybrid models, transfer learning strategies, ensemble techniques, etc., have achieved classification accuracies greater than 90%.

Furthermore, explainable AI (XAI) approaches, such as Grad-CAM, have improved interpretability and clinical trust, whereas lightweight architectures enable deployment in mobile and real-time healthcare settings. However, significant challenges persist despite these progressions.

Concerns of class imbalance, poor generalization to diverse datasets, and high computational overheads still persist and remain important obstacles to practical deployment.

Moreover, insufficient integration of complementary feature extraction techniques, a sparse level of interpretability, and insufficient translation of interpretation severely restrict their utility at clinical scales. To fill these gaps, we establish a computationally efficient.

Incorporating complementary feature extraction techniques and explainable AI methodologies will promote generalization, classification accuracy, and robustness, and facilitate early, reliable clinical diagnosis.

Related Work

Mampitiya et al. (2022) Based on the proposal of a low-computational-cost machine method of high accuracy, classification of skin cancer types for the medical image domain. Sinha et al. (2025) have enhanced it through attention mechanisms and feature visualization, thereby providing visual justifications for predictions that can enable dermatologists to make therapeutic decisions. Kumar et al. (2025) proposed a scalable, smart, and interpretable AI-based system for early-stage skin cancer detection, which can be integrated into web-based or mobile healthcare platforms for online clinical diagnosis.

Moreover, the Grad-CAM visualization helps in enhancing model interpretability, which is a significant factor for it to be a dependable candidate for clinical use. Kaminidevi et al. (2025) trained the algorithm on a labeled dataset and used methods such as transfer learning and data augmentation to overcome the issues of small medical image datasets. AI-driven dermatology tools can be used to help clinicians with diagnosis for the early phases of cancer, reducing manual workload through reduced workload and leading to better outcomes in patients, which the study demonstrates. Kumar et al. (2025): AI-based models not only offer the general features of AI-driven systems but also contribute substantially to clinical analysis and inform decision-making.

The proposed framework by Shamia et al. (2025) achieved remarkable enhancements, mainly due to its ability to address complex structures and imbalanced datasets, making it an excellent candidate for clinical application in dermatology. The framework that is proposed in Mehboob et al. (2025) is presented for evaluation based on ISIC data and achieves an even higher classification accuracy, 89%, compared with baselines trained on data until the end of October 2023. This extension also evaluates the application and usefulness of model adaptation to low-resource mobile applications.

The proposed model by Thilaka et al. (2025) developed the first such technique, which creates an economic and noninvasive method for skin disease diagnosis that can lead to early detection, diagnosis, and treatment. Patila et al. (2025) proposed deep learning methods to revolutionize

skin cancer diagnosis for a more extensive process of early detection, classification, and severity profiling for better patient care and life.

Interpretability, dataset diversity, and medical metadata may be at a premium to best optimize model end-results for better outcomes in subsequent works. The image segmentation is the first step in a study conducted by Jayasree et al. (2025); segmented images were obtained following the Orthogonal Least Squares (OLS) technique. Haque et al. (2025) reported model is competitive in terms of computational power with respect to well-known techniques like ResNet and EfficientNet, while the proposed model performs better with little processing resources.

The paper highlights the advantages of deep learning in the automatic identification of skin cancer while underscoring class imbalance and the requirement for improved generalization in the proposed model. Natha et al. (2024) achieved the highest mean F1-measure, recall, and precision results, and they suggested that Max Voting is the most robust and reliable method. The novelty of the present work is that skin cancer lesions are classified by Max Voting with more robustness and reliability. Shaheer et al. (2025) followed by VGG-19 and Mobilenet-v2 with 88%.

Finally, this chapter is informative of tradeoffs between metrics, and offers insights into the deployment of artificial intelligence-based skin cancer detection algorithms in practice. Khasanah et al. (2024) advocated that the findings of this study could help to better diagnose skin cancer.

Methodology

The skin cancer classification system described here is conceptualized as an organized, reproducible deep learning pipeline that incorporates cutting-edge feature extraction and new weighting techniques. The dataset used in this research is prepared from the ISIC (International Skin Imaging Collaboration) and Kaggle Skin Cancer: Malignant vs. Benign datasets, guaranteeing availability and integrity. The dataset comprises dermatoscopic images from two classes (benign and malignant), totaling N images. To remain consistent with pre-trained models, all images are resized to 224×224 pixels and normalized using the ImageNet mean and standard deviation.

The dataset is split into training and test sets at an 80:20 ratio to achieve balanced class distribution. All preprocessing is performed using the ImageFolder and DataLoader modules of PyTorch (batch size of 32 to improve efficiency). A pretrained ResNet-50 model is used as a backbone for feature extraction, with the final fully connected layer removed to obtain 2048-dimensional feature vectors.

The resulting features pass through a hybrid fully connected classifier with 2048, 512, and 128 neurons, followed by an output layer with the number of neurons first run in an array and trained with the Cross-Entropy loss for the number of classes, with an Optimizer for 50 epochs

and a learning rate of 0.001. One major novelty of this work is the inclusion of a feature-value-enhanced approach, the Gradient-Weighted Histogram Technique (GHT). GHT determines feature map gradients for class outputs, interprets absolute gradients into histograms, and assigns normalized importance scores for features.

Next, a GHT-weighted classification method (of the features) is employed to further train the classifier with a reduced learning rate of 0.0005 over 20 epochs, which enhances both the discriminative ability and interpretability. To maintain experimental reliability and reproducibility, we use various internal controls, such as a fixed random seed initialization (starting point), the same preprocessing steps (always the same), and standardized training and validation splits.

The model is assessed using gold-standard performance measures, such as accuracy, precision, recall, F1-score, and a confusion matrix, on the held-out test set. In addition, cross-validation (if performed) and comparison against baseline data models can be used to verify robustness and generalization performance. This approach is new in that it incorporates deep feature extraction and gradient-based feature importance weighting to improve binary (benign and malignant) classification performance at the edge while maintaining greater interpretability. GHT with a pretrained ResNet-50 backbone is an efficient and scalable framework suitable for real-world clinical application.

The resulting model architecture illustrates how to fully classify skin cancer. This is achieved through image preprocessing and dataset loading. Developers then perform deep feature extraction with the pretrained ResNet-50 network. They use this to build hybrid classifiers, refine features using the Gradient Histogram Technique (GHT), and evaluate performance with standard classification metrics.

Dataset Description

The skin cancer dataset is loaded from a specified directory containing the training images. Images are resized to 224×224 pixels to be compatible with the deep learning model. To balance computational efficiency and memory usage, a batch size of 32 was used during training. To speed up training, these experiments are run on the GPU when available; otherwise, on a CPU.

Table 1. The dataset used in this study consists of two classes: benign and malignant skin lesions. The training set includes 1,440 benign and 1,197 malignant images. The testing set consists of 360 benign images and 300 malignant images. The distribution of data samples ensures that both model training and evaluation requirements are met for both classes.

Image Transformation and Preprocessing

The skin cancer dataset is repeatedly transformed and preprocessed to achieve consistent results and improve

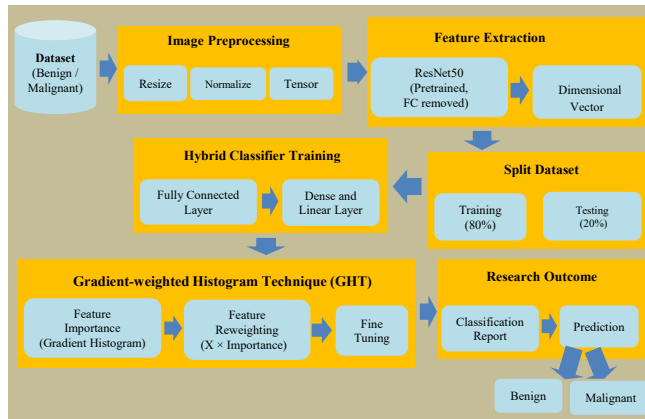


Figure 1: ResGrad Hybrid model of the proposed model

Table 1: Details of the Skin Cancer Dataset

Classes	Training dataset	Testing dataset
Benign	1440	360
Malignant	1197	300

model performance. First, images are resized to 224×224 pixels as required by ResNet-50. Next, these resized images are converted into PyTorch tensor format. Then, the tensors are normalized using ImageNet mean ([0.485, 0.456, 0.406]) and standard deviation ([0.229, 0.224, 0.225]) values. This sequence of resizing, tensor conversion, and normalization stabilizes training and accelerates convergence, enabling the pretrained ResNet-50 model to leverage its representation on the ImageNet dataset.

Display of Sample Images

The analysis of the skin cancer dataset begins with an initial visual assessment that illustrates how data preprocessing methods are implemented. A representative image is selected from the validation subset, with corresponding images and diagnostic labels clearly indicated. To process the image tensors, perform the following: first, convert them to

the height-width-channel (HWC) format with standardized spatial dimensions; then, restore their original appearance through ImageNet mean and standard deviation reverse normalization. The processed images are subsequently shown in a grid layout with class labels, omitting axis ticks to facilitate clearer visual interpretation. Finally, the visualization stage demonstrates dataset characteristics, verifies label accuracy, and allows a preliminary visual comparison between benign and malignant skin lesions.

Figure 2. Typical images of benign and malignant cases showing visual differences between benign and malignant samples.

ResGrad Hybrid Backbone

A pretrained ResNet-50 model serves as the basis for deep feature extraction in this study. The model is first initialized with ImageNet (IMAGENET1K_V2) pretrained weights, enabling successful transfer learning and faster training. The network needs a separate change: replace the fully connected final classification layer with an identity layer for feature extraction. We now use the model's feature representations, which replace the previous class-probability outputs. The system will transfer the backbone to the CPU or GPU for evaluation in a consistent manner, with feature extraction. ResNet-50 follows the same dual approach, generating strong 2048-dimensional features for the proposed, generating predictions to support future classifications.

Feature Extraction

The hybrid ResNet-50 backbone initialization yields deep features extracted from preprocessed skin cancer images. The images are processed through the ResNet-50 model in evaluation mode, which disables gradient computation to reduce processing requirements. The network generates high-level deep feature representations with 2048 features per input image batch. Feature vectors are reshaped to a two-dimensional format and aggregated from all batches.

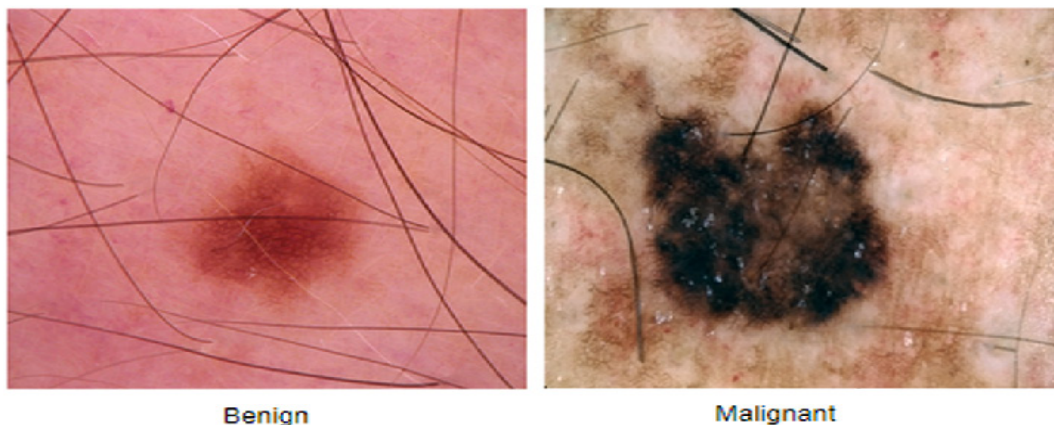


Figure 2: Images of Benign and Malignant cases

The features are transferred from the GPU to the CPU and converted into NumPy arrays for additional analysis. The system retains feature vectors and their class labels. A single feature matrix containing all extracted features serves as input to the upcoming classification process.

Train–Test Split

The dataset is first subjected to deep feature extraction. Following this, the data is split into training and test sets at an 80:20 ratio, with stratified sampling to evenly distribute benign and malignant cases. To ensure reproducibility, a fixed random seed is used. Next, the training and test sets, along with their labels, are converted to PyTorch tensors and transferred to the appropriate computational device. This preparation enables efficient model training, which is then used to assess performance during classification and fine-tuning.

ResGrad Hybrid Classifier

The classification of skin cancer uses a hybrid fully connected classifier based on extracted deep features. The classifier integrates multiple dense layers. The input layer accepts 2048-dimensional feature vectors from the ResNet-50 backbone. The system includes a fully connected layer with 512 neurons and ReLU activation, using a 0.4 dropout rate to control overfitting. It then applies a hidden layer with 128 neurons, ReLU activation, and a dropout rate of 0.3. The final output layer maps learned representations to the number of skin cancer classes. A Cross-Entropy loss function trains the classifier, using the Adam optimizer with a learning rate of 0.001. Training runs for 50 epochs, updating model parameters via backpropagation to reduce classification loss. The system logs training loss values after each epoch for convergence monitoring. This hybrid head enables effective learning from the deep feature representations extracted by the ResNet-50 backbone.

Result and Discussion

Based on the experimental evaluation, the baseline ResNet-50 achieved an accuracy of 87%, which is consistent with previous work showing that most standard transfer learning models (e.g., ResNet or VGG) have reported scores of 85–90% on benchmark datasets like ISIC. Using the proposed Gradient-Weighted Histogram Technique (GHT) and optimization approach, the final ResGrad Hybrid model achieves an accuracy of 89% and provides improved class-wise performance, as indicated by the confusion matrix and classification report, resulting in decreased misclassification of lesion categories. Compared with more recent state-of-the-art methods, higher accuracies have also been achieved with more complex architectures. Such work shows that performance can be significantly improved by increasing architectural complexity and using ensemble learning. Yet these high-performing models, too, inevitably entail a high

computational cost, larger parameter sizes, and a lower fitness for real-time or resource-limited clinical systems. Furthermore, most of those approaches are designed to maximize overall accuracy, but they neglect interpretation and feature-level understanding, which are very important in medical decision-making. Furthermore, dataset imbalance and poor model generalization remain issues of concern. In contrast, the proposed ResGrad Hybrid model achieves a good trade-off between performance, interpretability, and computational efficiency. The contribution of the work is primarily the implementation of the Gradient-Weighted Histogram Technique (GHT) to enhance discriminative feature learning by adding importance ratings based on graph-based gradient information. The proposed method instead aims to explicitly enhance feature representations, in addition to classical methods, to reduce confusion from visually similar lesions by achieving higher class-wise prediction. While the accuracy achieved (89%) is lower than that of highly complex or attention-based ensemble models, the proposed method enables several distinct advantages: (i) The proposed novel feature-weighting mechanism (GHT) improves both the interpretability and relevancy of features. (ii) Compared with a resource-consuming ensemble, only ResNet-50, which is a single backbone, reduces model complexity and makes it easier to interpret information. (iii) Improved binary (benign and malignant) classification capability with better class-wise discrimination. In conclusion, the results of this study show that GHT integration greatly improves the learning of discriminative features and, therefore, the proposed model can be considered a practical and scalable method for skin cancer classification. We show in this work that substantial performance improvements can be induced not only by complex architectural behavior but also by sophisticated feature-refinement schemes, and we fill in the gaps identified in previous work.

Standard Training

A hybrid classifier implements routine supervised learning techniques to evolve its training methodology. It trains for 50 epochs; i.e., its model runs in training mode to update some parameters. After the optimizer resets its gradients for predictions, the classifier receives the training feature set from each epoch, which helps produce forecasted class scores. The Cross-Entropy loss function computes classification loss based on how the predicted outputs match the real ground-truth labels. This involves backpropagating the gradients and updating the parameters with the Adam optimizer. To monitor the progress of model learning and its convergence pattern throughout the training period, the system stores training loss at each epoch.

Table 2. A summary of the training procedure that the model underwent in various training sessions is shown in

Table 2: Details of the Model Training Process Across Epochs

Epoch 1/50 Loss: 0.6909	Epoch 26/50 Loss: 0.2024
Epoch 2/50 Loss: 0.6460	Epoch 27/50 Loss: 0.1932
Epoch 3/50 Loss: 0.5908	Epoch 28/50 Loss: 0.1877
Epoch 4/50 Loss: 0.5318	Epoch 29/50 Loss: 0.1757
Epoch 5/50 Loss: 0.4778	Epoch 30/50 Loss: 0.1663
Epoch 6/50 Loss: 0.4384	Epoch 31/50 Loss: 0.1596
Epoch 7/50 Loss: 0.4138	Epoch 32/50 Loss: 0.1518
Epoch 8/50 Loss: 0.4015	Epoch 33/50 Loss: 0.1439
Epoch 9/50 Loss: 0.3785	Epoch 34/50 Loss: 0.1405
Epoch 10/50 Loss: 0.3602	Epoch 35/50 Loss: 0.1336
Epoch 11/50 Loss: 0.3414	Epoch 36/50 Loss: 0.1212
Epoch 12/50 Loss: 0.3298	Epoch 37/50 Loss: 0.1158
Epoch 13/50 Loss: 0.3179	Epoch 38/50 Loss: 0.1115
Epoch 14/50 Loss: 0.2985	Epoch 39/50 Loss: 0.1028
Epoch 15/50 Loss: 0.2975	Epoch 40/50 Loss: 0.0980
Epoch 16/50 Loss: 0.2780	Epoch 41/50 Loss: 0.0933
Epoch 17/50 Loss: 0.2726	Epoch 42/50 Loss: 0.0868
Epoch 18/50 Loss: 0.2618	Epoch 43/50 Loss: 0.0798
Epoch 19/50 Loss: 0.2572	Epoch 44/50 Loss: 0.0742
Epoch 20/50 Loss: 0.2484	Epoch 45/50 Loss: 0.0660
Epoch 21/50 Loss: 0.2329	Epoch 46/50 Loss: 0.0607
Epoch 22/50 Loss: 0.2367	Epoch 47/50 Loss: 0.0534
Epoch 23/50 Loss: 0.2235	Epoch 48/50 Loss: 0.0551
Epoch 24/50 Loss: 0.2152	Epoch 49/50 Loss: 0.0470
Epoch 25/50 Loss: 0.2058	Epoch 50/50 Loss: 0.0412

the table. The table shows the key performance metrics evaluated during training, including training and validation losses and training and validation accuracies. Evaluating these metrics during training allows checking the model's progress in learning and convergence.

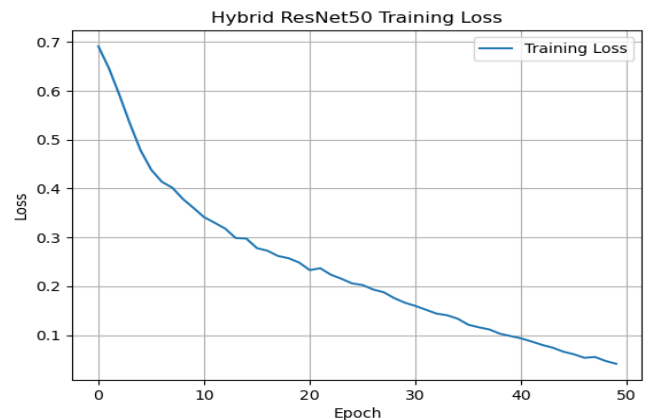
Loss Curve Analysis

The training loss reflects changes in the hybrid classification model's learning throughout all epochs. On this graph, the Cross-Entropy loss curve shows changes with each epoch. This pattern indicates progress toward the endpoint and overall stability. The training loss steadily declines, showing that the model learns efficiently and that the parameters are well-tuned. The Hybrid ResNet-50 model's loss curve helps identify underfitting or overfitting and confirms that the model converges during training.

Figure 3. The line chart shows the training loss of the Hybrid ResNet50 model over 50 training epochs. The vertical axis displays loss values which begin at approximately 0.7 and reach approximately 0.04 by the final epoch. The horizontal axis displays the duration of epochs, which extends from 0 to 50. The curve shows a downward trend, indicating that the model's training loss decreased throughout its training period as it acquired knowledge and reached learning milestones.

GHT Feature Weighting

The Gradient-Weighted Histogram Technique (GHT) estimates feature importance in several steps. First, the system evaluates the trained classifier to find the gradients of input features that influence specific class outputs. Next,

**Figure 3:** Training Loss of the ResGrad Hybrid Model

features from each class are selected, and the impact of their absolute gradient values on class prediction accuracy is tested. Histograms are then generated for each feature dimension, using fixed bins and scores weighted by gradient magnitude. These importance scores are normalized to ensure consistency. The normalized scores are applied element-wise to both the training

and test sets. This GHT-based approach highlights key features and reduces the impact of less relevant ones, thereby improving performance during classification and fine-tuning iterations.

GHT fine-tuning

The classifier is then adjusted using GHT-weighted features obtained in the previous stage of the GHT fine-tuning. An Adam optimizer with a reduced learning rate of 0.0005 is used to ensure stable, precise updates during training. For the model being trained, we use up to 20 epochs, during which GHT-weighted input features are used to compute gradients that update the parameters to minimize cross-entropy loss. At each epoch, the loss is recorded to monitor training progress. The fine-tuning allows the model to focus on discriminative features that the Gradient-weighted Histogram Technique can accentuate, improving the classification performance of the Hybrid ResNet-50 model and system reliability.

Table 3. The GHT model started to lose function, as its loss initially increased from 0.0417 to 0.2222 before it began to show consistent results. The training process showed continuous improvement, as the loss metric decreased throughout the session, indicating that the system learned to handle new incoming data. The final training epochs produced a loss statistic that stabilized in the range 0.032-0.035, indicating that the system had successfully reached its endpoint.

GHT loss curve

The GHT (Gradient-Weighted Histogram Technique) loss curves show the ResGrad Hybrid model's progress

Table 3: Overview of Model Training Progress During GHT Fine-Tuning Epochs			
Running GHT (Gradient-weighted Histogram Technique)...			
GHT Epoch 1/20 Loss: 0.0417	GHT Epoch 11/20 Loss: 0.0666		
GHT Epoch 2/20 Loss: 0.2222	GHT Epoch 12/20 Loss: 0.0422		
GHT Epoch 3/20 Loss: 0.0377	GHT Epoch 13/20 Loss: 0.0350		
GHT Epoch 4/20 Loss: 0.1185	GHT Epoch 14/20 Loss: 0.0427		
GHT Epoch 5/20 Loss: 0.1400	GHT Epoch 15/20 Loss: 0.0576		
GHT Epoch 6/20 Loss: 0.0623	GHT Epoch 16/20 Loss: 0.0579		
GHT Epoch 7/20 Loss: 0.0377	GHT Epoch 17/20 Loss: 0.0517		
GHT Epoch 8/20 Loss: 0.0501	GHT Epoch 18/20 Loss: 0.0384		
GHT Epoch 9/20 Loss: 0.0791	GHT Epoch 19/20 Loss: 0.0324		
GHT Epoch 10/20 Loss: 0.0871	GHT Epoch 20/20 Loss: 0.0349		

during GHT fine-tuning. In this step, the model adapts to task-specific data. The graph displays the convergence pattern and how the loss (error measure) changes over 20 fine-tuning epochs. (One epoch means a full pass through the training set.) The orange curve shows results for GHT-weighted feature optimization. This method gives more weight to important features and lowers classification loss. The model focuses on its most important features. The loss curve tracks the stability of GHT fine-tuning and training progress.

The graph shows initially high loss values that decrease continuously throughout the model's training. The curve stabilizes during the last training sessions, indicating that the model is approaching its final performance level.

Evaluation

The ResGrad Hybrid model, using GHT as its fine-tuning scheme, will undergo performance testing at the evaluation stage on the test dataset. After the model enters evaluation mode, GHT-weighted test features are used to make predictions. The sample's specific class with the highest predicted probability will be the final label. The results are then compared with ground truth labels to compute performance metrics - precision, recall, F1-score, and overall accuracy. The classification report also evaluates the model's overall performance at distinguishing between benign

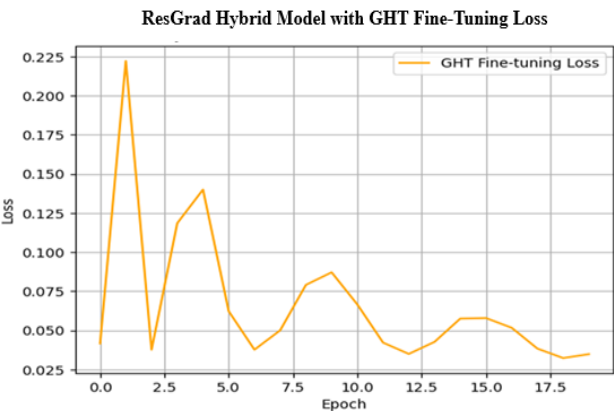


Figure 4: ResGrad Hybrid Model Fine-tuning Loss Curve

Table 4: Classification Results of the ResGrad Hybrid Model				
Precision	Recall	F1-score	Support	
benign	0.90	0.90	0.90	288
malignant	0.88	0.88	0.88	240
accuracy			0.89	528
macro avg	0.89	0.89	0.89	528
weighted avg	0.89	0.89	0.89	528

and malignant images across all skin cancer classes. This evaluation shows that GHT feature weighting, combined with the Hybrid ResNet-50 architecture, leads to better results.

Table 4. The ResGrad Hybrid model achieves an overall accuracy of 0.89 on the dataset. The system demonstrates excellent performance on both classes, with 0.90 precision, recall, and F1-score for the benign class and 0.88 for all three metrics for the malignant class. The system demonstrates equal performance across all classes, as both its macro and weighted averages reach 0.89.

Figure 5. The bar chart of performance metrics compares precision, recall, and F1-score for two different classes, benign and malignant, and the overall accuracy of the model. Since all three metrics are 0.90, precision, recall, and F1-score for the benign class are the same. The malignant class has slightly lower, but balanced values (precision, recall, and F1-score are all 0.88). The model achieves an overall accuracy of 0.89, demonstrating its ability to correctly classify both classes. The chart shows that the model performs consistently, with accuracy similar across all tests.

ResGrad Hybrid model confusion matrix

The ResGrad Hybrid model's confusion matrix visually summarizes its classification outcomes for the Benign and Malignant classes. It displays the counts of true positives, true negatives, false positives, and false negatives based on actual versus predicted labels from the GHT-weighted test dataset, using a blue color map. Each matrix cell details the number of predictions for each class pairing, enabling users to identify the frequency and type of errors. This

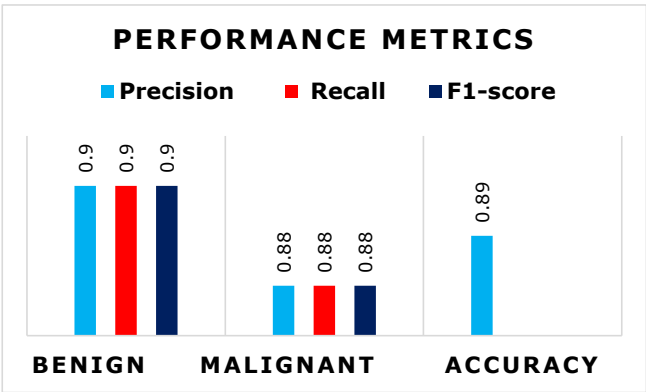


Figure 5: Evaluation Metrics for the Model

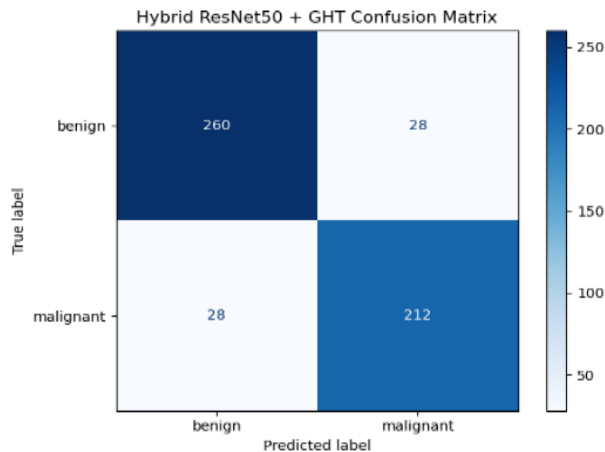


Figure 6: Confusion Matrix for the ResGrad Hybrid Model

visualization highlights how well the model distinguishes between benign and malignant skin lesions, clearly outlining its strengths and weaknesses.

Figure 6. The confusion matrix displays the classification results of the ResGrad Hybrid model on benign and malignant samples. The study achieved correct predictions for 260 of 288 benign cases, while 28 cases received incorrect malignant diagnoses. The model achieved balanced classification results, correctly diagnosing 212 malignant samples out of 240 total samples, while 28 samples received incorrect diagnoses as benign.

The deep learning model performance assessment in Table 5 demonstrates that the ResGrad Hybrid model achieves better results than all other tested methods. The system achieved its highest performance metrics: 89.4% accuracy, 90.28% precision, 90.28% recall, and 90.28% F1-score.

Figure 7 shows a bar chart of several convolutional neural network (CNN) architectures. Metrics include classification accuracy, precision, recall, and F1-score. These are based on model predictions and ground-truth labels. The ResGrad Hybrid model achieves the highest and most consistent results, reaching 89.4% for all metrics. This model outperforms all baselines, offering superior and more robust classification performance than other deep learning techniques.

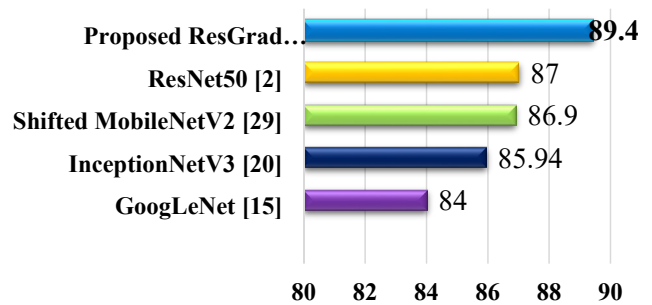


Figure 7: Performance Metrics of Comparison Models

Discussion

The experimental results show that the ResGrad Hybrid Model achieved an accuracy of 89.4%. It also had precision, recall, and F1-score values of 90.28%, which indicates a strong and reliable performance in identifying skin lesions. These findings support the idea that combining residual learning with gradient-based feature extraction improves classification and diagnostic accuracy. A comparison with other deep learning models indicates that the proposed model outperformed several earlier methods.

Mehra et al. (2021) used GoogLeNet for skin lesion classification and got an accuracy of 84.0%, which is notably lower than the proposed model's accuracy. While GoogLeNet effectively extracts features at multiple scales, its relatively shallow structure limits its ability to capture important lesion features.

Similarly, Noris et al. (2023) used InceptionNetV3 and reported an accuracy of 85.94%, a precision of 80.97%, and a recall of 86.89%. The ResGrad Hybrid Model exceeded these numbers by achieving higher precision and recall, showing improved sensitivity and fewer misclassifications. This enhancement comes from the better feature representation offered by residual learning and gradient-based optimization, which help in extracting complex lesion patterns.

Thurnhofer-Hemsi et al. (2021) introduced a Shifted MobileNetV2 model and reported an accuracy of 86.9%, with precision and recall values of 87.47% and 86.14%, respectively. Even though MobileNetV2 is efficient and

Table 5: Performance Comparison Between the Proposed Method and Existing Models

Deep learning model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
GoogLeNet (15)	84.00	--	--	--
InceptionNetV3 (20)	85.94	80.97	86.89	--
Shifted MobileNetV2 (29)	86.90	87.47	86.14	--
ResNet50 (2)	87.00	--	--	--
Proposed ResGrad Hybrid Model	89.4	90.28	90.28	90.28

light, the proposed model showed better performance, achieving higher accuracy and more balanced evaluation metrics. These results suggest that deeper residual feature learning can provide richer representations than lighter models, especially when dealing with subtle differences in skin lesion images.

Additionally, Bazgir et al. (2024) used ResNet50 and achieved an accuracy of 87.0%. Since ResNet50 already includes residual connections, its performance supports the effectiveness of residual learning for skin cancer classification. However, the ResGrad Hybrid Model raised the accuracy to 89.4%, indicating that adding gradient-based feature enhancement provides more discriminative information and leads to better generalization.

Overall, the comparative analysis confirms that the ResGrad Hybrid Model consistently outperforms the existing models in the literature. The higher accuracy, precision, recall, and F1-score show its strength and effectiveness for automated skin cancer detection.

These results agree with recent studies that highlight how hybrid and residual-based architectures can greatly improve classification performance by learning better features and reducing information loss during training [2,15,20,29]. Thus, the experimental findings support the proposed hypothesis and demonstrate the superiority of this framework over traditional deep learning methods for binary skin cancer classification.

Conclusion and Future Work

This research introduced a hybrid deep learning framework, the ResGrad Hybrid model, that combines a pretrained ResNet-50 backbone with the proposed GHT methods and employs the gradient-weighted histogram technique to improve skin cancer classification. This model is based on 3,297 images, comprising 2,637 training and 660 testing samples, including dermoscopic images and their corresponding evaluations. The results are presented quantitatively, with a summary of 89.4% accuracy, an increase from the baseline of 87% for the ResNet-50 approach, demonstrating a noticeable improvement of 2.4%. Besides the improvement in class-wise prediction, the reduction in misclassified data and the increased confusion matrix accuracy demonstrate that the proposed approach is a strong candidate. This work is at the forefront through the combination of gradient-based feature importance weighting (GHT) with deep feature extraction. In contrast to existing methods that rely solely on CNN-based representations or time-intensive ensemble methods, our approach provides a lightweight yet capable approach for estimating feature importance via gradient-derived histograms. Not only does this improve classification performance, but it also increases model interpretability, which is very important for clinical decision support systems. The proposed model addresses the important

problems in skin cancer classification - accurate binary (benign and malignant) prediction, lower dependency on manual diagnosis, and improved interpretability - thus closely mirroring the research objectives. Moreover, the model optimally balances performance and computational efficiency and applies to real-data clinical or telemedicine applications, given that timely and reliable diagnosis is critical. Although these results are encouraging, the study recognizes some limitations related to the dataset size, class imbalance, and generalization across diverse populations. Several potential actionable paths for future work are presented to tackle these limitations. To advance robustness and generalizability, the model can be expanded and improved on larger, more heterogeneous datasets containing other skin cancer subtypes first. Second, adopting better data augmentation techniques and cross-dataset validation strategies can improve the model's performance in the real world. Third, combining complementary Explainable AI (XAI) techniques with GHT could offer a more complete clinical understanding and enhance trust among healthcare practitioners. Finally, a lightweight deployment optimized for mobile or edge devices is also a route forward for high utilization in remote, resource-limited environments. Overall, this study presents evidence that deep learning with feature importance weighting is a promising approach towards better classification of skin cancer. The current ResGrad Hybrid model, proposed herein, provides a scalable, interpretable, and efficient approach and significantly contributes to the state of the art in AI-based computerized diagnostic systems.

Author Contributions

Conceptualization, S. V and Dr. J. G. R.S; methodology, S. V; software, S. V; validation, S. V and Dr. J. G. R. S; formal analysis, S. V; investigation, Dr. J. G. R. S; resources, Dr. J. G. R. S; data curation, S. V; writing original draft preparation, S. V; writing review and editing, Dr. J. G. R. S; visualization, S. V; supervision, Dr. J. G. R. S; project administration, Dr. J. G. R. S; funding acquisition, Dr. J. G. R. S. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

The author sincerely acknowledges the support and resources that enabled the development of the proposed deep learning-based skin cancer classification method. Special thanks are extended for contributions to dataset preparation, preprocessing, and implementation of the pretrained ResGrad Hybrid model using transfer learning techniques. The author also expresses gratitude to mentors and peers for their valuable guidance, suggestions, and continuous support throughout this work.

Ethical Statement

This study did not involve any experiments on humans or animals. Ethical approval is not applicable in this study.

Supplementary Information

This paper discusses a Hybrid ResNet-50 with the Gradient-weighted Histogram Technique for Automated Skin Cancer Classification. The proposed model improves upon the baseline ResNet-50 (87% accuracy) to the ResGrad Hybrid model (89.4%), demonstrating a notable advancement in automatic skin cancer detection performance.

References

- Akter, M. A., Khatun, R., Talukder, M. A., Islam, M. M., & Uddin, M. A. (2024). An integrated deep learning model for skin cancer detection using hybrid feature fusion technique [Preprint]. arXiv. <https://arxiv.org/abs/2410.14489>
- Bazgir, E., Haque, E., Maniruzzaman, M., & Hoque, R. (2024). Skin cancer classification using Inception network. *World Journal of Advanced Research and Reviews*, 21, 839–849. <https://doi.org/10.30574/wjarr.2024.21.2.0500>
- Fanconi, C. (2024). Skin cancer: Malignant vs. benign [Dataset]. Kaggle. <https://www.kaggle.com/datasets/fanconic/skin-cancer-malignant-vs-benign>
- Güler, F., & Agraz, M. (2025). Investigation of binary and multiclass classification performance of skin cancer images using transfer learning methods. *Journal of Clinical Practice Research*, 235–245. <https://doi.org/10.14744/cpr.2025.27623>
- Hadianti, S., & Kodri, W. A. G. (2023). Optimization of machine learning approach using Optuna in heart disease prediction. *Journal of Medical Informatics Technology*, 59–64. <https://doi.org/10.37034/medinftech.v1i3.15>
- Haque, M. I., Khatun, K., Rohman, G., Moon, M. S. A., & Billah, M. M. (2025). A deep learning framework for skin cancer diagnosis: Addressing class imbalance and performance optimization. *European Journal of Computer Science and Information Technology*, 13(1), 63–71. <https://doi.org/10.37745/ejcsit.2013/vol13n16371>
- Hasan, N., Nadaf, A., Imran, M., Jiba, U., & Sheikh, A. (2023). Skin cancer: Understanding the journey of transformation from conventional to advanced treatment approaches. *Molecular Cancer*, 1–70. <https://doi.org/10.1186/s12943-023-01854-3>
- Jayaseeli, J. D. D., Briskilal, J., Fancy, C., Vaitheeshwaran, V., Patibandla, & R. S. M. L. (2025). An intelligent framework for skin cancer detection and classification using fusion of squeeze-excitation DenseNet with metaheuristic-driven ensemble models. *Scientific Reports*, 1–33. <https://doi.org/10.1038/s41598-025-92293-1>
- Jayasree, M., Kumar, K. K., Chandrasekaran, G., Saranya, M., Gopinath, S., & Rajasekaran, T. (2025). WN-based skin cancer lesion segmentation of melanoma using deep learning methods. *Indonesian Journal of Electrical Engineering and Informatics*, 13, 451–463. <https://doi.org/10.52549/ijeei.v13i2.6151>
- Kaminidevi, S., Shanmugam, K., Kalaiselvi, R., Kurinjimalar, K., & Sathiya, S. (2025). Enhanced skin cancer detection and classification using convolutional neural network. *International Journal of Scientific Engineering Science*, 9, 73–80. <http://ijses.com/>
- Khasanah, N., & Winnarto, M. N. (2024). Application of deep learning with ResNet50 for early detection of melanoma skin cancer. *Journal of Medical Informatics Technology*, 2, 16–20. <https://doi.org/10.37034/medinftech.v2i1.31>
- Kumar, G. N. K., Koushika, B., Lahari, C., Sai, N. M., & Mahesh, P. (2025). Skin cancer detection using deep learning. *International Journal of Innovative Research in Science, Engineering and Technology*, 14, 8114–8121. <https://doi.org/10.1109/ICEARS53579.2022.9751826>
- Kumar, U. V., Nallamotheu, V. P., Yaswanth, T., & Pavanteja, P. (2025). Hybrid CNN-SVM pipeline for detecting skin cancer within lumps using transfer learning ResNet50 model. *Journal of Nonlinear Analysis and Optimization*, 16, 130–137. <https://doi.org/10.36893/JNAO.2025.V16I01.017>
- Mampitiya, L. I., Rathnayake, N., & De Silva, S. (2022). Efficient and low-cost skin cancer detection system implementation with a comparative study between traditional and CNN-based models. *Journal of Computational Cognitive Engineering*, 2(3), 226–235. <https://doi.org/10.47852/bonviewJCCE2202482>
- Mehra, A., Bhati, A., Kumar, A., & Malhotra, R. (2021). Skin cancer classification through transfer learning using ResNet-50. *Advances in Intelligent Systems and Computing*, 1300. https://doi.org/10.1007/978-981-33-4367-2_6
- Mehboob, S., Bukhari, M., Shah, Y. A., Khan, S., & Sharif, M. (2025). Enhanced skin cancer classification with MobileNetV3 and morphological preprocessing: A deep learning-based extension. *International Journal of Innovative Science and Technology*, 7, 1–12.
- Mijwil, M. M. (2021). Skin cancer disease images classification using deep learning solutions. *Multimedia Tools and Applications*, 80, 26255–26271. <https://doi.org/10.1007/s11042-021-10952-7>
- Nguyen, A. T. P., Jewel, R. M., & Akter, A. (2025). Comparative analysis of machine learning models for automated skin cancer detection: Advancements in diagnostic accuracy and AI integration. *American Journal of Medical Science and Pharmaceutical Research*, 7, 15–26. <https://doi.org/10.37547/tajmspr/Volume07Issue01-03>
- Natha, P., & RajaRajeswari, P. (2024). Advancing skin cancer prediction using ensemble models. *Computers*, 13, 1–24. <https://doi.org/10.3390/computers13070157>
- Noris, A. W. S. (2023). Implementasi deep learning untuk klasifikasi tanaman hias beracun menggunakan algoritma convolutional neural network, 6(1), 39–46. <https://sipora.polije.ac.id/id/eprint/19808>
- Patila, P. R., & Tandon, R. (2025). An efficient skin cancer segmentation and classification with severity analysis using HAM10000. *Journal of Neonatal Surgery*, 14, 1866–1877. <https://doi.org/10.52783/jns.v14.3050>
- Prome, J. N., Sultana, F., & Anika, S. (2024). Skin cancer detection: Leveraging hybrid deep learning models and traditional machine learning classifiers. *Proceedings of ICCA*, 770–778. <https://doi.org/10.1145/3723178.3723280>
- Ramadhan, A. A., Kareem, O. S., & Zeebaree, D. Q. (2025). A novel skin cancer detection approach using deep learning algorithm with image segmentation filters. *ARO Scientific Journal of Koya University*, 8, 153–161. <http://dx.doi.org/10.14500/aro.12024>
- Saputri, D. U. E., Khasanah, N., Aziz, F., & Hidayat, T. (2023). Enhancing skin cancer classification using optimized InceptionV3 model. *Journal of Medical Informatics Technology*, 65–69. <https://doi.org/10.37034/medinftech.v1i3.14>
- Shaheer, M., Arooj, T., Adeel, H., Saleem, T., & Rao, I. R. (2025). AI-based skin cancer detection algorithms: Opportunities, challenges and way forward. *Asian Journal of Science*,

- Engineering and Technology, 4, 87–109. <https://doi.org/10.47264/idea.ajset/4.1.6>
- Shamia, D., Umapriya, R., Prasad, M. L. M., Chowdhury, R., Kumar, P., & Vishnupriya, K. (2025). Enhancing skin cancer classification with mixup data augmentation and EfficientNet. *Journal of Electronics, Electromedical Engineering and Medical Informatics*, 7, 557–566. <https://doi.org/10.35882/jeeemi.v7i2.703>
- Sinha, S. S. R., & Sagar, A. K. (2025). Early stage skin cancer detection using deep learning: A comprehensive model for improved treatment outcomes and survival rates. *Journal of Information Systems Engineering and Management*, 10, 1052–1061. <https://doi.org/10.52783/jisem.v10i35s.6186>
- Thilaka, A., Nisha, V., & Nathiya, T. (2025). Leveraging SCNN for enhanced skin cancer detection using deep learning approaches. *Journal of Neonatal Surgery*, 14, 551–557. <https://www.jneonatsurg.com>
- Thurnhofer-Hemsi, E., López-Rubio, E., Domínguez, E., & Elizondo, D. A. (2021). Skin lesion classification by ensembles of deep convolutional networks and regularly spaced shifting. *IEEE Access*, 9, 112193–112205. <https://doi.org/10.1109/ACCESS.2021.3103410>