International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(6s) 2026 pp. 950-960
https://doi.org/10.70917/ijcisim-2026-2998
Article

Deep Learning-Based Skin Cancer Classification Using Hybrid S-ResNet and Explainable AI

C. Samdeepanmoses[1]*, P.S.Eliahim Jeevaraj [2]

[1*,2]Department of Computer Science, Bishop Heber College(Autonomous),
(Affiliated to Bharathidasan University, Trichy), Tiruchirappalli – 620017, India
sam.cs.res@bhc.edu.in[1*], eliahimps.cs@bhc.edu.in[2]
Corresponding Author: eliahimps.cs@bhc.edu.in[2]

Abstract:
Skin cancer is one of the most common and life-threatening diseases in the world. Early detection is essential in order to enhance treatment outcome and patient survival. In this work, a Hybrid S-ResNet framework with Grad-CAM is proposed for automated skin cancer classification using dermoscopic images. In the proposed approach, a modified architecture based on ResNet18 is used as a feature extractor to learn discriminative characteristics of lesions using residual learning. The input images are preprocessed using resizing, normalization and data augmentation techniques to improve the robustness of the model. Weighted Random Sampling, label smoothing and dropout regularization are employed during training to address class imbalance and improve generalization. The model is trained by AdamW optimizer with cosine annealing learning rate scheduling for better convergence and less overfitting. In addition, the proposed framework utilizes Gradient-weighted Class Activation Mapping (Grad-CAM) to produce visual explanations via highlighting the important parts of the image that lead to the prediction. The proposed model was compared with Resnet18, Resnet50, DenseNet121 and EfficientNetB0 under similar experiment settings. Experimental results show that the proposed Hybrid S-ResNet achieved better results with an accuracy of 84.90%, recall of 93.41%, F1-score of 90.69%, and AUC of 90.83%, proving the efficiency of the Hybrid S-ResNet to classify skin cancer. The combination of deep learning and explainable artificial intelligence offers a reliable framework to assist clinical decision making in dermatological diagnosis.
Keywords: Skin cancer, Hybrid S-ResNet, Grad-CAM, Deep learning, Explainable AI, ResNet18, AdamW optimizer.

1. Introduction

Skin cancer is one of the most common forms of cancer worldwide and continues to pose a significant public health challenge due to its increasing incidence and mortality rates. Among the various types of skin cancer, melanoma is considered the most aggressive because of its high metastatic potential and associated risk of death. Early detection plays a crucial role in improving treatment outcomes and patient survival. Conventional diagnostic methods primarily rely on dermoscopic examination and visual assessment using the ABCDE rule (Asymmetry, Border irregularity, Color variation, Diameter, and Evolution of skin lesions). However, these approaches are often subjective and may result in diagnostic inconsistencies. The fields of medical image analysis and computer-aided diagnosis (CAD) have shed their traditional reliance on manual labor by experiencing a significant shift toward automated solutions in recent years.In recent years, there have been major strides in the field of medical image analysis and computer-aided diagnosis (CAD) systems, which have been revolutionized by the advent of Artificial Intelligence (AI) and Deep Learning (DL). CNNs in particular have shown the potential of effectively learning useful semantic features from health-related images and accurately classifying them.. They have been used to help automatically detect and classify skin lesions, and have shown significant performance in learning complex visual patterns, which makes them promising tools for aiding dermatological diagnosis. Residual Networks (ResNet) is another popular CNN architecture that has been introduced due to the effectiveness of the residual learning mechanism for feature extraction and stable training. In the medical imaging domain, ResNet models have been reported to achieve good performance in various medical imaging tasks, such as skin cancer classification. Moreover, with the growing demand for systems that can be interpreted, Explainable Artificial Intelligence (XAI) has encouraged the use of techniques to share insights into the model decision making process and enhance the trust of the clinical environment. Motivated by these developments, this study proposes a Hybrid S-ResNet with Grad-CAM framework for automated skin cancer classification using dermoscopic images. The proposed approach aims to classify skin lesions into Malignant and Benign categories while providing visual explanations of the classification results. To assess its effectiveness, the proposed model is compared with established deep learning architectures, including ResNet18, ResNet50, DenseNet121, and EfficientNetB0. The study seeks to develop an accurate, computationally efficient, and interpretable framework that can support clinicians in the early detection of skin cancer.

Copyright:© 2026 by the authors

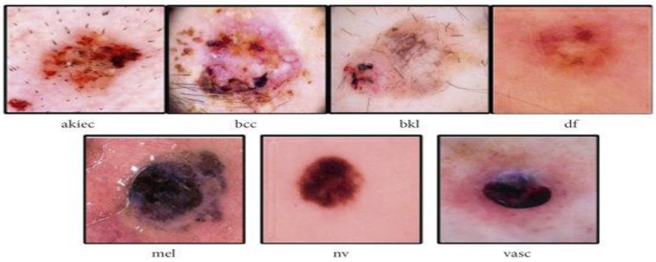

Fig. 1. Sample Skin cancer images

2. Existing works

Recent advances in deep learning have significantly improved the performance of automated skin cancer diagnosis systems. Numerous studies have explored the use of Convolutional Neural Networks (CNNs), transfer learning techniques, and hybrid architectures for the classification of dermoscopic skin lesion images. Pretrained models such as ResNet, DenseNet, EfficientNet, and VGG have been widely adopted due to their ability to extract discriminative features from complex medical images. In addition, researchers have investigated various optimization strategies, data augmentation techniques, and class-balancing methods to enhance classification accuracy and model generalization. More recently, Explainable Artificial Intelligence (XAI) approaches such as Grad-CAM have been incorporated into deep learning frameworks to improve model interpretability and clinical trust. This section reviews the existing literature on skin cancer classification, highlighting the strengths and limitations of previous approaches and establishing the motivation for the proposed Hybrid S-ResNet with Grad-CAM framework.

Jain *et al.* proposed a hybrid deep learning model for skin cancer detection that integrates ResNet50 for feature extraction with Support Vector Machines (SVMs) for classification. This approach leverages ResNet50's ability to automatically learn hierarchical features from dermoscopic images, enhancing the model's capacity to distinguish between malignant and benign lesions. By combining these features with SVMs, known for their effectiveness in both binary and multiclass classification tasks, the model achieves improved diagnostic accuracy. The study demonstrates that this hybrid model not only enhances classification performance but also offers a robust and scalable solution for skin cancer detection. The authors highlight the potential of this approach to assist clinicians in early diagnosis, thereby improving patient outcomes. This work contributes to the growing body of research aimed at integrating deep learning techniques into medical diagnostics, particularly in dermatology[1].

Kumar *et al.* explored the application of linear and radial basis function (RBF) kernels in Support Vector Machines (SVMs) for skin cancer detection. They observed that linear SVMs perform effectively when the data is linearly separable, offering simplicity and computational efficiency. However, for more complex datasets with non-linear relationships, RBF kernels provide a significant advantage by mapping data into higher-dimensional spaces, allowing for better separation and classification. The authors highlighted that selecting the appropriate kernel is crucial for enhancing model performance, especially in medical imaging tasks where data patterns can be intricate. Their findings underscore the importance of kernel choice in developing robust deep learning models for accurate skin cancer detection. By tailoring the kernel to the dataset's characteristics, clinicians can achieve more reliable and precise diagnostic outcomes. This research contributes to the ongoing efforts to integrate advanced machine learning techniques into dermatological diagnostics[2].

Shaik *et al.* emphasized that early detection and prompt medical treatment significantly enhance survival rates, especially for melanoma. They noted that dermatologists typically rely on dermoscopic images to visually examine patients for skin cancer diagnosis, with biopsies sometimes required for confirmation. However, the authors highlighted that manual diagnosis carries inherent disadvantages, such as subjectivity and inter-observer variability, which can lead to inconsistencies in diagnosis. To address these challenges, they proposed integrating machine learning algorithms with Principal Component Analysis (PCA) and Gabor filters for feature extraction. This approach aims to automate the diagnostic process, reduce human error, and improve the accuracy and efficiency of skin cancer detection. By leveraging advanced computational techniques, the study contributes to the development of more reliable and scalable diagnostic tools in dermatology. Such advancements are crucial for enhancing early detection and treatment outcomes for skin cancer patients[3].

A physician's ability to check beneath the skin using specialized light and oil determines how Bhatt et al. proposed an automated system for lung nodule classification that integrates ResNet50 for feature extraction with Support Vector Machines (SVMs) for classification. This hybrid approach leverages ResNet50's deep convolutional layers to capture complex patterns in medical images, enhancing the model's ability to distinguish between benign and malignant nodules. By employing SVMs, known for their effectiveness in high-dimensional spaces, the system achieves improved classification accuracy. The authors noted that this model offers a robust and scalable solution for lung cancer detection, potentially aiding clinicians in early diagnosis. The study contributes to the growing body of research aimed at integrating deep learning techniques into medical diagnostics, particularly in oncology. Such advancements are crucial for enhancing early detection and treatment outcomes for cancer patients. The findings underscore the potential of combining deep learning and machine learning methods to develop effective diagnostic tools [4].

Iourzikene, Gougam *et al.* developed an automated system for brain tumor detection using MRI images, emphasizing the role of artificial intelligence (AI) in enhancing diagnostic accuracy. The system comprises three stages: pre-processing to enhance image quality, feature extraction using techniques like Bag of Features (BoF) and ResNet50, and classification via Support Vector Machines (SVMs) with linear, quadratic, and cubic kernels. The authors highlighted that AI, particularly through the emulation of human-like cognitive processes, enables machines to perform tasks such as pattern recognition and logical reasoning, which are crucial for accurate medical diagnostics. Their findings demonstrated that the BoF-SVM combination achieved a 100% classification accuracy, outperforming the ResNet50-SVM configurations, which achieved accuracies ranging from 96.2% to 98.6%. This underscores the potential of AI in automating complex diagnostic tasks, reducing human error, and facilitating early detection of brain tumors. The study contributes to the growing body of research integrating AI into medical imaging, aiming to improve patient outcomes through timely and precise diagnoses[5].

Gaffoor et al. proposed an effective approach for skin disease detection and classification by integrating ResNet-50 for feature extraction with Support Vector Machines (SVMs) for classification. This hybrid model leverages ResNet-50's deep convolutional layers to automatically learn hierarchical features from dermoscopic images, enhancing the model's capacity to distinguish between various skin conditions. By employing SVMs, known for their effectiveness in high-dimensional spaces, the system achieves improved classification accuracy. The authors noted that while deep learning models like ResNet-50 offer high accuracy, they often require significant computational resources and time for training. To address this, they emphasized the importance of suitable feature selection techniques to reduce the dimensionality of the data, thereby decreasing training time and improving prediction speed without compromising accuracy. The study suggests that future research should focus on developing customized convolutional neural network (CNN) models tailored for skin disease classification and exploring more effective feature selection methods to enhance model efficiency and performance. This approach aims to make dermatological diagnosis more accessible and efficient, particularly in resource-limited settings [6].

Chatrapathy *et al.* emphasized the importance of effective data preprocessing for skin cancer classification using a hybrid CNN-SVM model. They highlighted *data normalization* as a critical step, which scales input data to a specific range, such as [0,1] or [−1,1], to reduce singularities, mitigate the effects of outliers, and improve model stability. Common approaches include min-max normalization, z-score normalization, and neural network normalization, with z-score normalization particularly effective in algorithms relying on distance-based similarity measures. The authors also underscored the role of *data augmentation* in increasing dataset diversity and enhancing model generalization, achieved by introducing minor variations or generating synthetic samples from existing images. These techniques collectively improve the robustness, accuracy, and efficiency of deep learning models, enabling more reliable classification of skin lesions [7].

Saeed *et al.* explored methods to provide domain-specific explanations for convolutional neural network (CNN) models in the context of skin cancer classification. They emphasized that while single-sample explanations are useful for understanding individual predictions, gaining insight into the model's overall behavior across multiple cases is equally important. The authors proposed an explanation framework that aggregates insights from numerous samples, allowing clinicians to identify consistent patterns and features that the CNN considers most relevant. This approach not only improves interpretability but also helps validate the reliability of the model by revealing potential biases or misclassifications. Saeed *et al.* highlighted that domain-specific explanations can guide clinicians in assessing whether the model focuses on clinically meaningful aspects of lesions, thereby increasing trust in AI-assisted diagnosis. Furthermore, combining multiple explanations can provide a comprehensive understanding of feature importance and model decision-making strategies across different lesion types. The study concluded that such interpretability methods are crucial for integrating deep learning models into clinical workflows, enhancing both transparency and the practical utility of AI in dermatology[8].

Musthafa *et al.* introduced an optimized Convolutional Neural Network (CNN) architecture to enhance the accuracy and efficiency of skin cancer diagnosis. Utilizing the HAM10000 dataset, which comprises a diverse range of dermatoscopic images, the researchers developed a sophisticated CNN model designed to capture intricate visual features of skin lesions. To address class imbalance within the dataset, they implemented an innovative data augmentation strategy, ensuring a balanced representation of each lesion category during training. The model's architecture includes multiple convolutional, pooling, and dense layers, aimed at effectively learning from the complex image data. Training was optimized using the Adam optimizer, fine-tuned over 50 epochs with a batch size of 128, and incorporated a Model Checkpoint callback to preserve the best model iteration. The proposed model achieved an impressive accuracy of 97.78%, with precision and recall both at 97.9%, and an F2 score of 97.8%, demonstrating its potential as a robust tool in the early detection and classification of skin cancer. These advancements underscore the significant role of deep learning in automating dermatological diagnostics and improving patient outcomes [9].

Zareen *et al.* introduced a hybrid deep learning model that combines Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to improve skin cancer diagnosis. They utilized ResNet-50 as a feature extractor, enabling the model to capture complex spatial hierarchies within dermoscopic images. The incorporation of RNNs allowed the system to model sequential and contextual dependencies, enhancing its ability to differentiate between subtle lesion patterns. The model was evaluated on a dataset of 9,000 images covering nine different skin cancer types. The results demonstrated high recognition accuracy, highlighting the robustness and effectiveness of the hybrid approach. By combining CNNs and RNNs, the model leverages both spatial and sequential information, which is critical for accurate medical image analysis. This study emphasizes the potential of hybrid architectures in developing reliable, AI-driven diagnostic tools for early and precise skin cancer detection [10].

Vischer *et al.* compared multiple convolutional neural network (CNN) architectures, including ResNet50, DenseNet201, and EfficientNet, for feature extraction in classifying histopathology images of skin diseases. The evaluation was conducted on a dataset containing 11 distinct skin disease classes. ResNet50 demonstrated robust performance across most categories, highlighting its ability to effectively capture complex image features. DenseNet201 and EfficientNet also showed strong results, but ResNet50's residual connections facilitated better gradient flow and efficient learning in deep networks. The study emphasized the importance of selecting an appropriate CNN backbone to optimize performance for specific medical imaging tasks. ResNet50's combination of depth, accuracy, and computational efficiency makes it particularly suitable for dermatological image classification. These findings provide valuable guidance for developing reliable and automated diagnostic systems in clinical dermatology[11].

Md Sirajul Islam *et al.* explored the application of transfer learning techniques for binary classification of skin cancer images, focusing on distinguishing between benign and malignant lesions. Utilizing the publicly available ISIC dataset, they fine-tuned five pre-trained models—ResNet-50, MobileNet, InceptionV3, DenseNet-169, and InceptionResNetV2—by adjusting various layers and activation functions to enhance model accuracy. To improve model stability and generalization, data augmentation techniques were applied, introducing variations such as rotations and flips to the training images. The experimental setup involved evaluating different hyperparameters, including batch sizes, epochs, and optimizers, to identify the optimal configuration for each model. Among the models tested, ResNet-50 achieved the highest performance, attaining an accuracy of 93.5%, precision of 94%, recall of 77%, F1-score of 86%, and an ROC-AUC score of 86.1%. This study demonstrates the effectiveness of transfer learning in skin cancer classification tasks, highlighting the potential of fine-tuning pre-trained models to achieve high accuracy with limited datasets. The findings suggest that employing transfer learning, coupled with data augmentation, can significantly enhance the performance of deep learning models in medical image classification tasks [12].

Samira Mavaddati *et al.* introduced a hybrid deep learning model that integrates ResNet50 with Long Short-Term Memory (LSTM) networks to enhance skin cancer classification. This approach leverages ResNet50's powerful feature extraction capabilities alongside LSTM's proficiency in capturing sequential dependencies, enabling the model to better understand the structural content of skin lesions. The hybrid model, termed ResNet50-LSTM, utilizes transfer learning techniques, starting with a pre-trained ResNet50 model and fine-tuning it for the specific task of skin cancer classification. By combining the strengths of both deep learning architectures, the model aims to improve accuracy and reduce false positives in skin cancer detection. This innovative approach addresses the limitations of traditional deep learning models by incorporating sequential learning, which is particularly beneficial for analyzing the complex textures and patterns in skin lesion images. The study's findings suggest that the ResNet50-LSTM hybrid model offers a promising direction for developing more accurate and reliable automated systems for skin cancer diagnosis [13]. In order to categorize segmented photos of skin lesions, the study assesses the effectiveness of Convolutional Neural Networks (CNN) and ResNet50. To improve the quality of the input data, the Kaggle dataset was pre-processed using U-Net for lesion segmentation. Accuracy, precision, recall, and F1-score metrics were used to train and assess both models [14].

3. Methodology

The architecture of the Hybrid S-ResNet with Grad-CAM model to give a clear understanding of the proposed skin cancer classification framework is shown in Fig. 2. The first step in the framework is the acquisition of dermoscopic images of the skin lesions from the HAM10000 dataset and then the preprocessing operations. The images are resized to 224 × 224 size and then the images are normalized with ImageNet mean and std values for compatibility with the pretrained network. Data augmentation, such as random horizontal flipping, random rotation and color jittering, is used during training to increase model robustness and generalization ability. The skin lesion datasets frequently have a class imbalance problem, so a Weighted Random Sampler is applied for a balance learning of cancerous and non-cancerous samples.

The preprocessed images are then fed into a pretrained ResNet18 back bone with weights obtained from the ImageNet dataset. The architecture for the proposed S-ResNet is based on multiple stages of convolutional and residual blocks (Conv2_x–Conv5_x) to learn discriminative spatial features from dermoscopic images. Instead of ResNet18's original classification layer, a customized classification head comprised of Global Average Pooling, a fully connected layer (512 → 128), ReLU activation, dropout regularization, and a binary output layer is used. This light-weight residual learning framework makes it less complex without compromising its ability to extract features for skin cancer classification. Then the feature representation extracted is classified as Malignant and Benign.

Transfer learning uses pretrained knowledge from large-scale picture datasets to improve learning performance. The suggested model is trained with weight decay regularization and the AdamW optimizer at a learning rate of 1×10^{-2}. Additionally, to enhance generalization and lessen prediction overconfidence, Cross-Entropy Loss with label smoothing is applied. To improve convergence and training stability, a Cosine Annealing Learning Rate Scheduler is integrated to dynamically modify the learning rate during training. Together, these optimization techniques improve classification performance on unknown samples and lessen overfitting.

The framework incorporates Gradient-weighted Class Activation Mapping (Grad-CAM) as an Explainable Artificial Intelligence (XAI) method to enhance model interpretability. The Grad-CAM module produces visual heatmaps that show which areas of the image have the greatest influence on the categorization result. Clinicians can confirm if the model concentrates on medically relevant lesion locations during prediction by superimposing these activation maps onto the original dermoscopic images. Finally, the effectiveness of the proposed framework is assessed through comparative analysis with ResNet18, ResNet50, DenseNet121, and EfficientNetB0 using performance metrics including Accuracy, Precision, Recall, F1-Score, AUC, Specificity, Sensitivity, and Confusion Matrix analysis [15].

3.1 ResNet18

ResNet18 is a lightweight deep convolutional neural network that employs residual learning to overcome the degradation problem commonly observed in deep architectures. Data can move directly between the 18 layers of the model thanks to shortcut or skip connections. During training, these residual connections assist effective feature learning and enhance gradient propagation. ResNet18 is appropriate for applications with constrained computer resources because of its lower computational complexity and fewer parameters. The architecture starts with a convolution layer and a max-pooling layer to get the low level image features. A number of RBDs are then trained to progressively learn feature representations from the input images. Residual blocks learn feature transformations while at the same time learning to retain the original information with identity mappings. This mechanism helps to prevent vanishing gradient issue and ensures stable learning behavior within the network. The simplicity and power of its architecture has made ResNet18 well suited for a variety of image classification tasks. The model is used widely in medical image analysis due to the trade-off between computational efficiency and the model's predictive power. The architecture has the potential to learn texture, shape and color variations in dermoscopic images in skin lesion analysis. The residual learning strategy enables the network to retain a crucial set of lesion properties when extracting features. Based on this, ResNet18 was used as a baseline model to assess the effectiveness of using Residual feature extraction for skin cancer image classification.

3.2 ResNet50

The ResNet50 is another deeper convolutional neural network with 50 layers, built on the principles of residual learning. The architecture is based on bottleneck residual blocks, which enable learning rich feature representations while at the same time being computationally efficient. An efficient feature extraction is realized using the 1×1, 3×3 and 1×1 convolution operations in the bottleneck structure in order to reduce the number of parameters. The model's depth allows it to better detect and recognize subtle visual patterns and complex lesion features in dermoscopic images. Residual connections can help mitigate vanishing gradient and enhance stability of training in deeper networks. The network learns hierarchical representations of images from low (edges) to high (semantic) levels of discrimination. ResNet50 has the ability to learn complex lesion patterns that might not be represented by smaller networks. The architecture has been proven to excel at a range of computer vision tasks such as object recognition and medical diagnosis tasks. The model has a larger depth which enables it to learn richer contextual information from the input image. But deeper architectures may need more computational resources, and more training times. ResNet50 is used to classify skin cancer, and it can detect subtle differences in the structures and textures of lesions. The detailed features captured by the model help to enhance classification accuracy. In this paper ResNet50 was employed to study the effects of deeper hierarchical feature extraction on the classification performance of skin cancer.

3.3 DenseNet121

The DenseNet121 architecture is a densely connected CNN consisting of 121 layers and each layer is fed with feature maps from all the previous layers. This dense connectivity mechanism boosts feature reuse and boosts information flow across the network. DenseNet is different from the traditional architecture in that it makes direct connections between layers, allowing features to be shared throughout the network. This way, feature learning is not repeated and training gradient propagation is enhanced. Unlike the architecture of residual networks, the architecture of this architecture does not add up elements of the previous layers, but instead concatenates them. This enables later layers to have the capability of accessing both low-level and high level representations at the same time. DenseNet121 also outperforms the deep architectures with fewer parameters for classification tasks. Efficient reuse of features helps to achieve higher computational efficiency and lower model complexity. Dense connections help alleviate the vanishing gradient problem and improve convergence during optimization. The architecture has been proven to work well in image classification and medical image analysis applications. For image classification, DenseNet121 can learn a wide variety of characteristics of lesions and image patterns. The model is able to use the previously gained knowledge in an efficient manner, to enhance the prediction capability. For this study, DenseNet121 is added to explore the merits of dense feature connectivity in skin lesion classification.

3.4 EfficientNetB0

EfficientNetB0 is a deep learning architecture created by a compound scaling method that combines network depth, width and input image resolution parameters. The model scales more than one network dimension uniformly to maximize performance rather than just scaling one dimension. By utilizing this strategy, EfficientNetB0 can be trained to achieve better classification accuracy while using fewer parameters and computational resources. Neural architecture search techniques were employed to find an optimal architecture. EfficientNetB0 is a compound scaling deep learning architecture that scales the network depth, width and input image resolution simultaneously. The model scales several dimensions uniformly rather than scaling one dimension up, to optimize the performance of the model.

EfficientNetB0 uses mobile inverted bottleneck convolution (MBConv) layers, which add to its efficiency in learning features. Squeeze-and-excitation mechanisms are also used to enhance the feature representation for each channel. Architecture provides good tradeoff between computational efficiency and prediction performance. EfficientNetB0 achieves competitive accuracy using a fraction of the model parameters of conventional models. This means that the model can be used in applications that have limited hardware resources. The network has been shown to perform well on several image recognition tasks. EfficientNetB0 has the capability to learn discriminative visual features from complex medical data sets in medical image analysis. It is very efficient, and can be trained and inferred faster than others without compromising on accuracy. In this paper, EfficientNetB0 has been used as an example model for its benefits in terms of computational efficiency and classification capability in skin cancer analysis, which has lesser number of parameters and with less computational cost. The architecture was designed using neural architecture search techniques to learn an optimal architecture for the network.

3.5 Hybrid S-ResNet (Proposed Model)

The proposed Hybrid S-ResNet model is derived from a modified ResNet18 network architecture, combined with a custom classification network and a customized explainability module Grad-CAM. The architecture is specifically designed for classification of skin cancer with ensuring the interpretability of the prediction outcomes. Just like residual networks, the model uses skip connections to ensure smooth information flow between different layers. The framework starts with the dermoscopic image pre-processing, in which the images are resized and normalized for a standard input representation. Multiple residual block groups are used for feature extraction, where each group learns discriminative lesion properties. These layers retain low level image information like edges and textures and slowly build up a high level representation of the image content. Global Average Pooling is performed after the feature extraction step to reduce the size of the features without losing significant information. The extracted feature vectors are then forwarded to fully connected layers for classification. ReLU activation functions are employed to add non-linearity to learn better. To avoid overfitting and to enhance the generalization performance, dropout regularization is used. The network is trained to predict categories of lesions (cancer or not cancer) from learned feature representations. Additionally, the integrated Grad-CAM module generates visual heatmaps that identify important image regions contributing to predictions. These heatmaps provide explainability by highlighting lesion areas that influence model decisions. Such interpretability is important in medical applications where transparent decision-making is required. The performance and effectiveness of the proposed model are evaluated through comparative analysis with existing deep learning architectures under identical experimental settings.

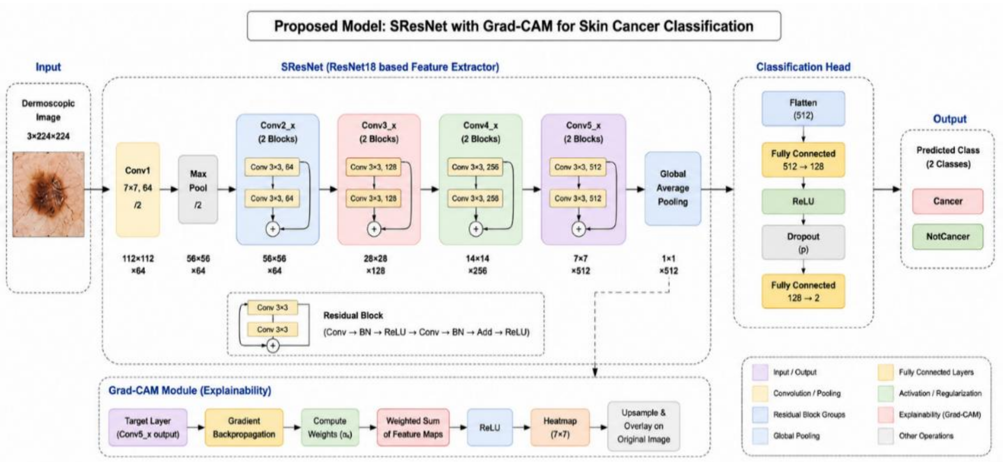

Fig.2. Flow of the proposed model

accuracy, particularly in cases where traditional neural networks may struggle. This hybrid approach results in a robust and efficient image classification model.

Algorithm : Proposed Hybrid S-ResNet for Skin Cancer Classification

Input: Binary skin lesion images (*Cancer, NotCancer*)

Output: Predicted labels, performance metrics, and Grad-CAM visualizations

1. Load dermoscopic images and corresponding labels from the dataset.
2. Split the dataset into training and testing sets using stratified partitioning.
3. Resize all images to 224 × 224 × 3 and normalize using ImageNet values.

4. Apply data augmentation using random flip, rotation, and color jitter.
5. Handle class imbalance using Weighted Random Sampling.
6. Initialize pretrained ResNet18 with ImageNet weights.
7. Remove the original classification layer from ResNet18.
8. Add Adaptive Average Pooling, Fully Connected, ReLU, and Dropout layers.
9. Construct the proposed Hybrid S-ResNet for binary classification.
10. Train the model using AdamW optimizer, label smoothing, and cosine annealing scheduler.
11. Evaluate performance using Accuracy, Precision, Recall, F1-score, and Confusion Matrix.
12. Generate Grad-CAM heatmaps for visual interpretation of predictions.

The proposed algorithm introduces an S-ResNet-based framework for binary skin cancer classification integrated with explainable AI techniques. Initially, the skin lesion dataset undergoes preprocessing and augmentation to improve generalization and reduce overfitting. The pretrained ResNet18 architecture is employed as a feature extraction backbone and enhanced with custom classification layers for improved learning performance. Hyperparameter optimization is performed to determine optimal training parameters, while an ablation study using different train–test splits and epoch configurations analyzes model convergence and robustness. Finally, Grad-CAM is applied to generate visual explanations, highlighting important lesion regions and improving the interpretability of the classification results.

4. Results and Discussion

Using transfer learning and EAI techniques, the proposed Hybrid S-ResNet-based framework was implemented for binary skin cancer classification. The data set was divided into two classes: Cancer and NotCancer, and processed by image augmentation and normalization techniques to enhance models generalization. To examine the learning behavior and convergence of the model, multiple train–test split configurations (80:20, 70:30, 60:40) and different epoch numbers (25, 50, 75, and 100) were used. A weighted random sampling strategy was used during training to overcome the data imbalance. The experimental results show that the proposed approach successfully learned discriminative features from skin lesion images and gave robust classification performance under various experimental settings.

Hyperparameter optimizations were conducted to find appropriate training configurations with Optuna. The critical parameters that were optimized during the process are: learning rate, weight decay and dropout rate. Within the trials that were evaluated, the best parameters that were found were a learning rate of about 0.000252, a weight decay of 1.99×10^{-6}, and a dropout rate of 0.379. The performance of the model was sensitive to the selection of the parameters, as the accuracy values of the validation results were widely different in each trial during the optimization process. The highest accuracy for the best trial was around 84.61%, and there were a number of trials which yielded lower accuracy because it contained improperly combined regularization and learning rate values. This observation reiterates the need to systematically tune the hyper parameters to gain improved model performance and better convergence.

Investigating training and testing performance over several epochs gave insights on the learning behavior of the proposed architecture. The training accuracy is consistently improved with increasing number of epochs because the network is able to learn more complex representations from the dataset with more epochs. It was noted, however, that the training accuracy increased more rapidly than the testing accuracy, thus suggesting overfitting. The model was trained with dropout and label smoothing but the performance difference between training and testing performance was observed at higher epoch numbers. This indicates that the model gradually began to memorize the training samples instead of learning the general pattern. This kind of behavior could be typical when training a deep learning model with a small dataset and could suggest that it would be worthwhile to add more data to the training set or to use more powerful regularization methods to improve the generalization power of the model.

The ablation study with varying train-test ratio and epoch settings helped gain deeper insights into model robustness and convergence. The model learned better with a larger proportion of training data, as there were more samples of the data presented during training. The 80:20 split resulted in more training data and convergence properties than the 60:40 split. Likewise, the more the epochs, the better the features were learned at first, but too many epochs also led to overfitting. Thus, the ablation analysis revealed that a good balance between training data size and training time is key to optimal model performance. This study has confirmed the effectiveness of the proposed framework and also the factors that affect the framework's performance.

The framework is supplemented with explainable AI (XAI) analysis based on Grad-CAM, to enhance its interpretability and transparency. Grad-CAM computed activation maps by highlighting areas of the images of skin lesions that most influenced the classification decision. The visualization was created not only from prediction scores, but from showing regions of lesions, and creating contour overlays for better localization. The resulting activation maps showed that the model concentrated on relevant lesion regions and not on random background regions, enhancing the trustworthiness and interpretability of the model. The visual explanations in medical applications are especially useful, as they help the clinician comprehend the prediction of the model, and they also help give added confidence to a diagnostic automated system. Overall, the proposed S-ResNet framework improved the explainability by incorporating Grad-CAM, while maintaining the classification accuracy.

4.1 Ablation Study Results for Skin Cancer Classification

This section presents the ablation study conducted to analyze the contribution of different components in the proposed Hybrid S-ResNet framework and to identify the most effective architecture for skin cancer classification. The study was performed on the HAM10000 dataset by progressively incorporating various techniques including data augmentation, weighted sampling, label smoothing, and scheduler with dropout regularization. Initially, a baseline ResNet18 model was used, followed by the sequential addition of these components to evaluate their impact on model performance. The input images were resized and normalized to ensure compatibility with the network architecture. Data augmentation techniques improved model robustness, while the Weighted Random Sampler addressed class imbalance during training. Furthermore, label smoothing and Cosine Annealing scheduling with dropout were incorporated to enhance generalization and reduce overfitting. Each component was selected based on preliminary experiments and existing literature that highlighted their influence on deep learning model performance [13].

Performance Metrics:

The amount of accurate forecasting the model made out of all the samples is known as accuracy. By dividing the number of accurate predictions by the total number of samples evaluated, it displays the model's overall correctness[10].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Precision, also known as Positive Predictive Value, quantifies the proportion of samples that are truly correct when projected to be positive. It demonstrates how well the model predicts favorable outcomes[10].

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

Among the original sample's positive samples, recall is the likelihood of accurately predicting a positive sample.

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

F1-score, Recall and precision are incompatible metrics. Generally speaking, recall is frequently poor when precision is high; conversely, recall is frequently high when precision is low[10].

$$F1 = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (4)$$

In medical research, particularly the categorization of skin cancer, the kappa coefficient (κ) is a statistical measure that is frequently used to assess the degree of agreement between two raters or classification systems that goes above what would be predicted by chance. It is computed as:

$$\kappa = \frac{Po-Pe}{1-Pe} \quad (5)$$

The Geometric Mean (G-Mean) is a performance metric used to evaluate imbalanced classification problems. It balances sensitivity (recall) and specificity to ensure that both classes are well predicted.

$$\text{G-Mean} = \sqrt{\text{Sensitivity} \times \text{Specificity}} \quad (6)$$

S.No	Model	Training Accuracy	Testing Accuracy
1	**ResNet18**	0.9138	0.8301
2	**ResNet50**	0.9150	0.8396
3	**DenseNet121**	0.9740	0.8301
4	**EfficientNetB0**	0.9115	0.8113
5	**Hybrid S-ResNet**	0.9929	0.8632

Table 1. Training and Testing Accuracy Comparison of Deep Learning Models for Skin Cancer Classification

S. No	Model	Precision	Recall	F1-Score	AUC
1	**ResNet18**	0.8659	0.9281	0.8959	0.8357
2	**ResNet50**	0.8982	0.8982	0.8982	0.8650
3	**DenseNet121**	0.8786	0.9101	0.8941	0.8786

4	**EfficientNetB0**	0.8713	0.8922	0.8816	0.8552
5	**Hybrid S-ResNet**	0.8813	0.9341	0.9069	0.9083

Table 2. Comparative Performance Evaluation of Deep Learning Models Using Precision, Recall, F1-Score, and AUC

The comparative performance analysis of different deep learning models reveals notable variations in their ability to classify skin cancer lesions accurately. Tables 1 and 2 summarize the training performance, testing performance, and evaluation metrics obtained by ResNet18, ResNet50, DenseNet121, EfficientNetB0, and the proposed Hybrid S-ResNet model. As presented in Table 1, the proposed Hybrid S-ResNet achieved the highest training accuracy of 99.29% and testing accuracy of 86.32%, indicating superior learning capability and generalization performance compared to the other evaluated architectures. ResNet50 obtained the second-highest testing accuracy of 83.96%, while ResNet18 and DenseNet121 achieved identical testing accuracies of 83.01%. EfficientNetB0 recorded the lowest testing accuracy of 81.13%. The evaluation metrics reported in Table 2 further demonstrate the effectiveness of the proposed model. Hybrid S-ResNet achieved the highest Recall (93.41%), F1-Score (90.69%), and AUC (0.9083), indicating a stronger ability to correctly identify cancerous lesions while maintaining balanced classification performance. Although ResNet50 achieved the highest Precision (89.82%), the proposed model maintained competitive precision (88.13%) while outperforming all other models in overall predictive capability. DenseNet121 and ResNet18 produced comparable results across most performance metrics, whereas EfficientNetB0 showed relatively lower classification effectiveness. Overall, the findings indicate that the proposed Hybrid S-ResNet framework provides the most reliable and robust performance for binary skin cancer classification among the evaluated deep learning models.

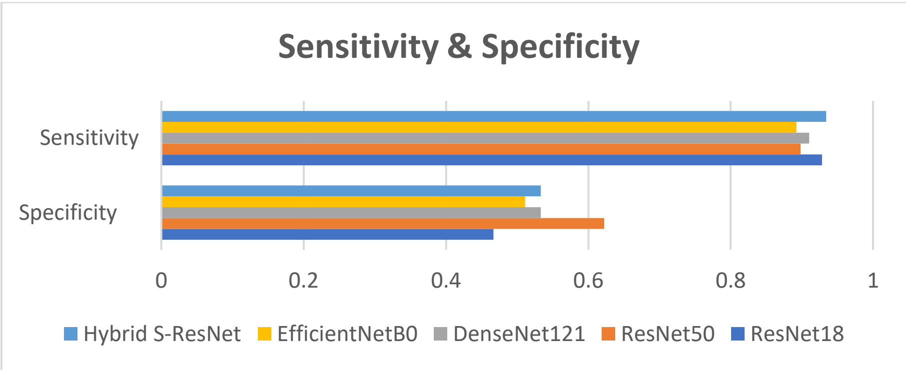

Fig.3 Sensitivity & Specificity

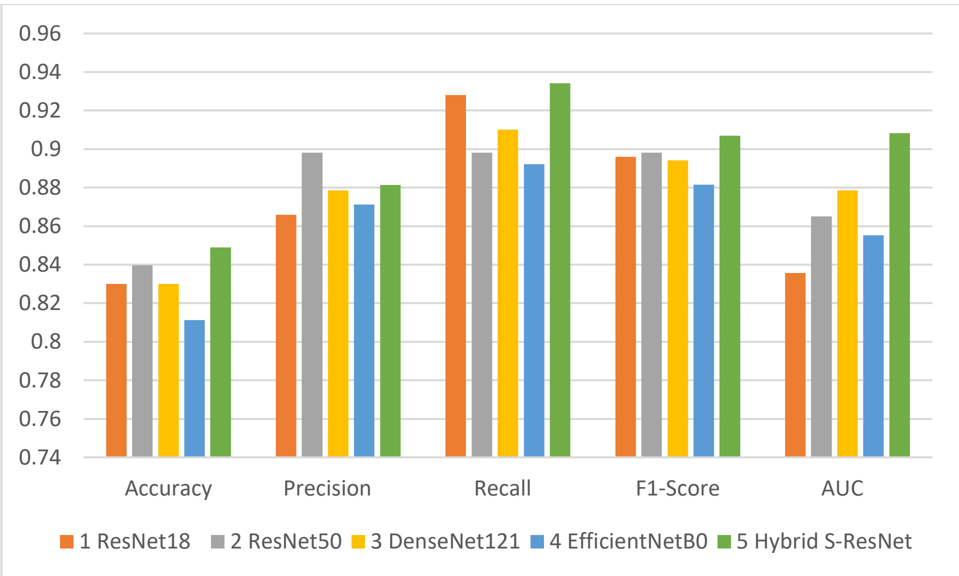

Fig.4 Evaluation Metrics Comparison for Skin Cancer Classification Models

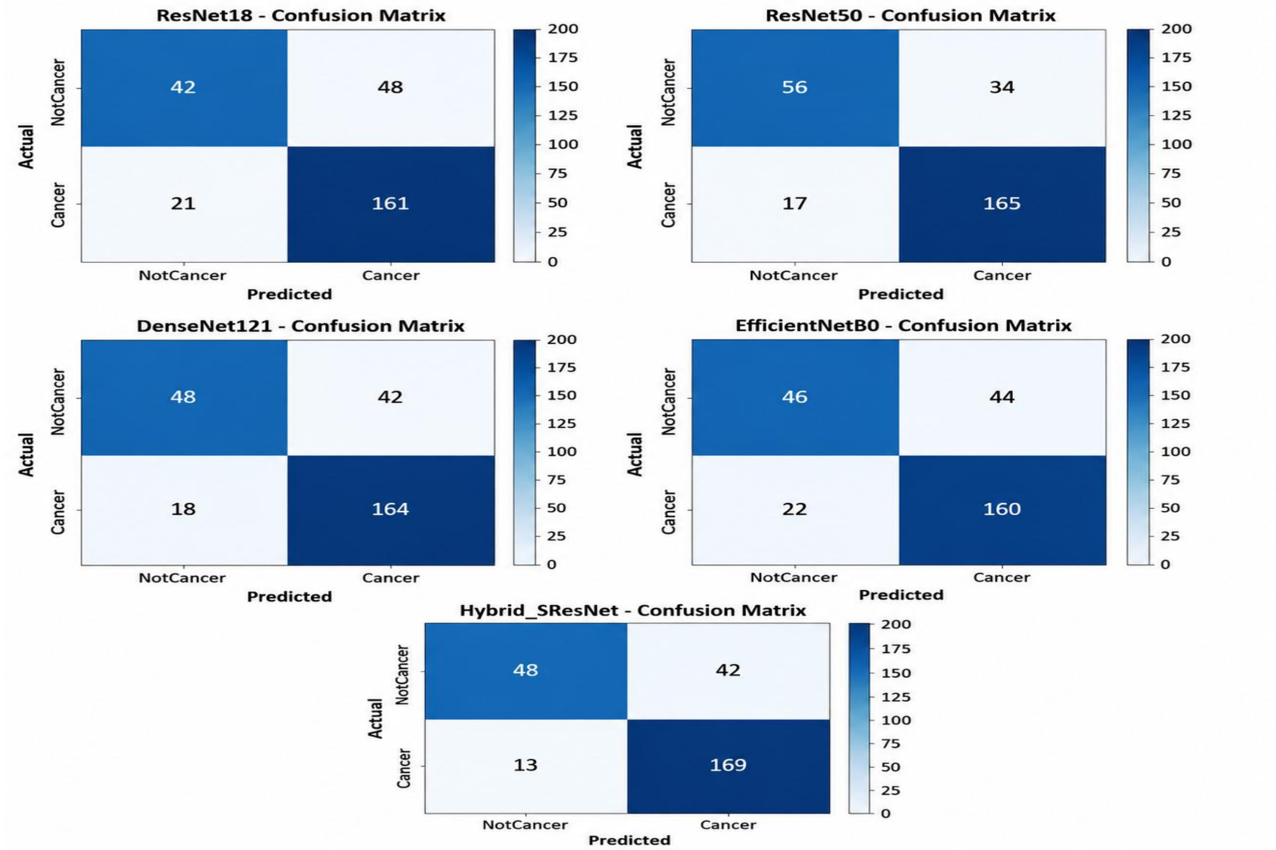

Fig. 5. Comparative Diagnostic Performance of Deep Learning Models for Skin Cancer Classification

Figure 5 illustrates the confusion matrix comparison of ResNet18, ResNet50, DenseNet121, EfficientNetB0, and the proposed Hybrid S-ResNet for skin cancer classification. The confusion matrices provide a detailed representation of correctly and incorrectly classified samples for both Malignant and Benign categories. Among the evaluated models, the proposed Hybrid S-ResNet demonstrated improved classification capability by correctly identifying 169 cancer cases while producing only 13 false negative predictions, indicating stronger sensitivity toward cancer detection. ResNet50 also exhibited competitive performance; however, the proposed model achieved a better balance between correct classification and error reduction. In contrast, ResNet18, DenseNet121, and EfficientNetB0 showed relatively higher misclassification rates. The results suggest that the proposed Hybrid S-ResNet effectively learned discriminative lesion characteristics and improved classification performance in distinguishing cancerous and non-cancerous skin lesions. While the surrounding ones show the number of samples that were mistakenly classified, the diagonal of the matrix shows the number of true positives. Xin *et al.* [16] proposed an improved transformer-based network for skin cancer classification using dermoscopic images. The model utilized self-attention mechanisms to capture discriminative lesion features more effectively than traditional CNNs. Experimental results demonstrated improved classification performance and robustness for skin lesion diagnosis. Toprak *et al.* [17] developed a hybrid convolutional neural network model for multi-class skin cancer classification. The proposed architecture combined multiple CNN components to enhance feature extraction and classification accuracy. Their results showed that hybrid CNN models can improve the reliability of automated skin cancer diagnosis. Bechelli *et al.* [18] conducted a comparative study of machine learning and deep learning algorithms for skin cancer classification from dermoscopic images. The study highlighted the superior performance of deep learning models in extracting complex lesion features. The authors emphasized the importance of selecting appropriate architectures to achieve accurate and reliable skin cancer diagnosis.

5. Conclusion

The proposed study presented a Hybrid S-ResNet with Grad-CAM framework for automated skin cancer classification using dermoscopic images. The model utilized a modified ResNet18-based architecture with transfer learning to extract meaningful and discriminative lesion features for binary classification into Cancer and NotCancer categories. To improve model robustness and generalization, preprocessing techniques including image normalization, data augmentation, and Weighted Random Sampling were incorporated during training. In addition, label smoothing, dropout regularization, and the AdamW optimizer with cosine annealing learning-rate scheduling were employed to enhance convergence and minimize overfitting. Experimental results demonstrated that the proposed Hybrid S-ResNet achieved improved performance compared to baseline models such as ResNet18, ResNet50, DenseNet121, and EfficientNetB0, attaining an accuracy of 84.90% and strong classification capability. Furthermore, the integration of Grad-CAM provided visual explanations by highlighting significant image regions contributing to prediction outcomes, thereby improving model interpretability and transparency. The findings indicate that the proposed framework can support reliable and explainable skin cancer diagnosis. Future work may focus on expanding dataset diversity, further optimizing model parameters, and exploring advanced explainable deep learning architectures for improved clinical applicability.

Acknowledgment

The authors express their sincere gratitude to the Management of Bishop Heber College, the University Grants Commission (UGC) - College of Excellence, and the Department of Science and Technology (DST) - FIST for their invaluable support in facilitating this research work."

References

1. Sweta Jain; Srushti Dhakate; Prathamesh Gujar; Prathamesh Rajbhoj, "Skin Cancer Detection and Classification using Deep Learning" *IEEE Xplore International Conference on Electronics and Renewable Systems (ICEARS)* ISBN:978-1-6654-8426-8 (2024)
2. R. Senthil Kumar, Amarjeet Singh, Sparsha Srinath, Nimal Kurien Thomas, Vishal Arasu "Skin Cancer Detection using Deep Learning" IEEE *Xplore* ISBN:979-8-3503-4986-3 (2023)
3. Abdul Rahaman Shaik, P. Rajesh Kumar, "Performance Evaluation of Machine Learning Algorithms on Skin Cancer Data Set Using Principal Component Analysis and Gabor Filters" IAENG International Journal of Computer Science Volume 51, Issue 7, July 2024, Pages 831-841 https://www.iaeng.org/IJCS/issues_v51/issue_7/IJCS_51_7_13.pdf
4. Shital D. Bhatt; Himanshu B. Soni; Heena R. Kher; Tanmay D. Pawar, "Automated System for Lung Nodule Classification Based on ResNet50 and SVM", IEEE *Xplore* 2022 3rd International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), ISBN: 978-1-6654-8269-1 (2023) Automated System for Lung Nodule Classification Based on ResNet50 and SVM | IEEE Conference Publication | IEEE Xplore
5. Zouhir Iourzikene, Fawzi Gougam, Djamel Benazzouz, "Performance Evaluation of Feature Extraction and SVM for Brain Tumor Detection Using MRI Images" Performance Evaluation of Feature Extraction and SVM for Brain Tumor Detection Using MRI Images | IIETA
6. Noorunisa Gaffoor; Sarmad Soomro,"Skin Disease Detection and Classification Using ResNet-50 and Support Vector Machine: An Effective Approach for Dermatological Diagnosis" 2023 IEEE International Conference on Internet of Things and Intelligence Systems (IoTaIS) https://ieeexplore.ieee.org/document/10346059
7. Chatrapathy K; M Lakshmi Prasad; Ajmeera Kiran; Pundru Chandra Shaker Reddy; G Charles Babu; N Partheeban,"Skin Cancer Classification using A Hybrid Convolutional Neural Network with SVM Classifier" https://ieeexplore.ieee.org/document/10426054
8. Mudqssir Saeed , Asma Naseer, Hassan Masoodi, Shafiq ur Rehman,"The Power of Generative AI to Augment for Enhanced Skin Cancer Classification: A Deep Learning Approach" Digital ObjectIdentifier10.1109/ACCESS.2023.3332628https://ieeexplore.ieee.org/abstract/document/10318035
9. M Mohamed Musthafa, Mahesh T R, Vinoth Kumar V & Suresh Guluwadi "Enhanced skin cancer diagnosis using optimized CNN architecture and checkpoints for automated dermatological lesion classification" (2024)
10. https://bmcmedimaging.biomedcentral.com/articles/10.1186/s12880-024-01356-8?utm_source=chatgpt.com#citeas
11. Syeda Shamaila Zareen, Guangmin Sun, Mahwish Kundi, Syed Furqan Qadri, Salman Qadri "Enhancing Skin Cancer Diagnosis with Deep Learning: A Hybrid CNN-RNN Approach" (2024) https://doi.org/10.32604/cmc.2024.047418
12. Anna-Lena Vischer, Jiayu Liu, Sinclair Rockwell-Kollmann, Stefan Günther & Klemens Schnattinger, "Skin Cancer Classification: A Comparison of CNN-Backbones for Feature-Extraction | SpringerLink" Skin Cancer Classification: A Comparison of CNN-Backbones for Feature-Extraction | SpringerLink
13. Md Sirajul Islam, Sanjeev Panta, "Skin Cancer Images Classification using Transfer Learning Techniques" https://doi.org/10.48550/arXiv.2406.12954
14. Samira Mavaddati, "Skin cancer classification based on a hybrid deep model and long short-term memory" https://doi.org/10.1016/j.bspc.2024.107109
15. Aris Wahyu Murdiyanto, Dian Hafidh Zulfikar, Bagus Satrio Waluyo Poetro, Alda Cendekia Siregar, "Performance Comparison of CNN and ResNet50 for Skin Cancer Classification Using U-Net Segmented Images" https://doi.org/10.56705/ijodas.v5i3.200
16. Jaisakthi S M, Mirunalini P, Chandrabose Aravindan, Rajagopal Appavu ,"Classification of skin cancer from dermoscopic images using deep neural network architectures" Multimedia Tools and Applications (2023) 82:15763–15778 https://doi.org/10.1007/s11042-022-13847-3
17. Chao Xin a,1, Zhifang Liu a,1, Keyu Zhao a, Linlin Miao a, "An improved transformer network for skin cancer classification" Computers in Biology and Medicine 149 (2022) 105939 https://www.sciencedirect.com/science/article/pii/S0010482522006746
18. Ahmet Nusret Toprak, Ibrahim Aruk, "A Hybrid Convolutional Neural Network Model for the Classification of Multi-Class Skin Cancer", International Journal of Imaging Systems and Technology https://onlinelibrary.wiley.com/doi/epdf/10.1002/ima.23180
19. Solene Bechelli 1,2 and Jerome Delhommelle "Machine Learning and Deep Learning Algorithms for Skin Cancer Classification from Dermoscopic Images" Bioengineering 2022, 9, 97. https://doi.org/10.3390/bioengineering9030097
20. Abdullah Al Mahmud, Sami Azam, Inam Ullah Khan, and Sidratul Montaha,"SkinNet-14: a deep learning framework for accurate skin cancer classification using low-resolution dermoscopy images with optimized training time" Neural Computing and Applications (2024) 36:18935–18959 https://doi.org/10.1007/s00521-024-10225-y
21. Rukhsar Sabir & Tahir Mehmood,"Classification of melanoma skin Cancer based on Image Data Set using different neural networks" Scientific Reports (2024) 14:29704 https://doi.org/10. 1038/s41598-024-75143-4