Data, Information and Computing Science
A. Nayyar et al. (Eds.)

doi:10.3233/ATDE250013

Deep Learning for Early Skin Cancer Detection: A Comparative Study on Hybrid CNN Models

Sakshi MATIRAY[1] and Dr. Law Kumar SINGH[2]
[1]Computer Engineering and Applications, GLA University, Mathura, India
[2]Amity School of Engineering and Technology, Amity University Punjab Mohali
[1] Sakshimatiray@gmail.com; [2] lawkumarcs1@gmail.com

Abstract: Globally, skin cancer—especially melanoma—is a major health concern, and better patient outcomes depend on early and precise diagnosis. Manual examination is a common component of traditional diagnostic techniques, but it is labor-intensive and subject to error. Promising approaches to automate skin lesion analysis have been made possible by recent developments in deep learning. In order to overcome the shortcomings of current approaches, this work investigates the creation of a hybrid deep learning model that blends VGG16 and ResNet50 architectures. The suggested approach incorporates sophisticated preprocessing methods, such as picture normalization and data augmentation, to improve feature extraction and classification accuracy using datasets like ISIC 2018 and HAM10000.Important results show that the hybrid model outperforms state-of-the-art benchmarks in performance metrics, achieving 98.75% training accuracy and 97.50% validation accuracy. The model's enhanced precision (97.60%), recall (97.55%), and F1 score (97.58%) highlight how reliable it is at differentiating between benign and malignant tumors. The study also emphasizes how crucial it is to strike a balance between model complexity and computing performance in order to support practical clinical deployment. This study advances computer-aided diagnostic (CAD) systems for the identification of skin cancer by tackling issues including class imbalance and dataset unpredictability. The suggested method has a great deal of promise to improve early diagnosis, lessen the need for invasive treatments, and assist dermatologists in making clinical decisions. This paper thoroughly reviews current methodologies used in skin cancer detection, offering a timely resource for researchers developing automated and precise melanoma detection models.

Keywords. CNN, Skin cancer, Melanoma, non-melanoma, Deep learning, Machine Learning, VGG16, ResNet50, EfficientNetV2-M

1. Introduction

Skin disorders pose major health risks to people and are frequently underdiagnosed, particularly in their early stages. It's a common misperception that conditions like acne or rashes will resolve themselves. Skin disorders pose major health risks to people and are frequently underdiagnosed, particularly in their early stages. It's a common misperception that conditions like acne or rashes will resolve themselves, and the World Health Organization has reported a marked increase in the prevalence of skin cancers, encompassing both melanoma and non-melanoma varieties. However, ignoring early indicators can lead to more serious complications, including skin cancer, A major contributing element to this trend is prolonged exposure to the sun's damaging UV rays, which can result in long-term skin damage (American Cancer Society, 2023). Skin health is also impacted by poor dietary choices, such as eating processed meals and produce that has been chemically treated. Many people usually overlook early preventive steps, enabling skin diseases to deteriorate over time, especially those with little health awareness. Environmental pollution and poor skincare habits lead to long-term harm that frequently goes unnoticed until it has greatly advanced. the World Health Organization notes a steady rise in both non-melanoma and melanoma skin cancer incidence globally. Extended exposure to the sun's harmful UV rays, which can result in long-term skin damage, is a significant contributing element to this trend. Dietary habits are just as significant as environmental influences. Increased consumption of processed foods and produce grown with excessive amounts of chemicals can hurt skin health. Many people, especially those with low health awareness, frequently neglect early preventive measures, allowing skin conditions to worsen over time. Poor skincare practices and environmental pollution cause long-term damage that frequently goes undetected until it has greatly advanced.

Furthermore, skin conditions have significant psychological effects. Many people suffer from a loss of confidence, which affects their social relationships and sense of self. Skin conditions have well-established psychological effects, such as a reduction in self-esteem and social anxiety. Small rashes or pimples that start as minor skin imperfections can develop into more severe and noticeable conditions, which exacerbate emotional and mental suffering. Ignoring these early warning indicators can result in issues that get harder to handle. Therefore, raising awareness, encouraging early identification, and offering efficient treatment are essential for dealing with skin illnesses, which may not seem like much at first but can develop into major health problems if prompt medical attention is not received. It's critical to diagnose skin cancer early, for avoiding problems and enhancing patient outcomes. By implementing preventive measures, such as regular skin exams and careful dietary choices, the likelihood of developing serious skin illnesses can be significantly decreased. In recent years, a variety of datasets containing dermoscopy images of skin lesions have advanced computer-aided diagnosis (CAD) systems for the analysis of skin cancer. This study addresses the shortcomings of conventional diagnostic techniques by utilizing deep learning techniques to enhance the early identification and precise categorization of skin malignancies. Our goal is to increase the overall effectiveness and precision of skin cancer diagnosis by utilizing a hybrid model that combines the advantages of VGG16 and ResNet50.

The structure of the paper is as follows: A thorough analysis of current methods for detecting skin cancer is given in Section 2, along with an emphasis on the main obstacles and developments in the field. The datasets used in this investigation are covered in Section 3, along with their attributes and applicability to model evaluation. The methodology is explained in Section 4, with particular attention paid to the experimental setup, preprocessing techniques, and hybrid model design. The results are examined in Section 5, which highlights the performance advantages of the suggested strategy by contrasting it with cutting-edge methods. A review of the results, their applications, and possible avenues for further study are included in Section 6's conclusion. The Conclusion of Future Directions is presented in Section 7, and the references for all of the papers are displayed in Section 8.

1.1 Motivation and Objectives:

Melanoma is the primary cause of skin cancer-related deaths, making it one of the most prevalent and deadly types of cancer in the world. Even though machine learning and deep learning have made great strides in early detection, issues like class imbalance, lesion feature unpredictability, and dataset restrictions frequently cause present models to perform poorly. These deficiencies highlight how urgently strong, precise, and scalable diagnostic instruments are needed to help physicians detect melanoma in a timely and trustworthy manner. This study's main goal is to improve skin cancer early detection and classification by tackling these issues using a hybrid deep-learning approach. The goal of this study is to create a model with better accuracy, precision, and recall by utilizing the complementing capabilities of VGG16 and ResNet50. The study's objective is to evaluate and contrast cutting-edge deep learning methods for classifying skin lesions. To create and assess a CNN hybrid model that combines the benefits of ResNet50 and VGG16. To evaluate the suggested model's performance using important metrics like accuracy, F1 score, and sensitivity while testing it on reputable datasets like HAM10000 and ISIC. To shed light on the suggested model's real-world use in clinical contexts, tackling issues including dataset bias and computational expenses.

2. Literature Review

The goal of Ghosh et al. [1] We used a dataset of 3000 photos from nine different skin diseases in our investigation. We used class weights throughout training to address the problem of class imbalance. As part of our methodology, we combined the architectures of VGG16 and ResNet50 to create a hybrid DL model. The hybrid model was trained for 100 epochs, and our results showed that it achieved 98.75% training accuracy and 97.50% validation accuracy, with strong performance metrics including a precision of 97.60%, recall of 97.55%, and an F1 score of 97.58%. These results confirm the potential of advanced DL models in the accurate detection and classification of skin cancer. Going forward, future work will focus on improving our model’s dependability in real-world clinical environments and addressing issues like robustness and generalization across diverse populations. As suggested by Vipin Venugopa et al. [2]in this paper, I used several datasets, such as ISIC 2018, ISIC 2019, and ISIC 2020. A total of 58,032 refined dermoscopic pictures were divided into seven classes, and HAM10000. I used deep learning architectures to improve model performance, specifically EfficientNet-B4 (Tan Le), and EfficientNetV2-M. Additionally, I employed techniques like transfer learning and data augmentation. Performance was evaluated using parameters like memory, accuracy, precision, and F1 score on both binary and multiclass classification tasks. The model's astounding 99.23% accuracy rate for binary classification and 97.62% for multiclass classification on the ISIC 2020 dataset. Transfer learning is applied to fine-tune the model, enabling classification into malignant (e.g., melanoma) and benign categories with features extracted from dermoscopic images. Walaa Gouda et al. (2022)[3], proposed The work made use of the 3533-image ISIC2018 dataset, which was enriched and enhanced with ESRGAN to increase diversity. To determine if skin lesions were benign or malignant, a CNN model and transfer learning architectures like ResNet50, InceptionV3, and Inception ResNet were used. Preprocessing included contrast enhancement, scaling, and normalization. InceptionV3, the top-performing model, with an accuracy of 85.7%, which is on par with dermatologist-level competence.

In order to strengthen diagnostic skills, future research recommends testing on bigger datasets and investigating architectures like DenseNet and VGG. Kinnor Das et al.(2021)[4] investigate the use of machine learning (ML) in skin cancer diagnosis, emphasizing key datasets, methodologies, architectures, techniques, performance, and outcomes. Two primary datasets are used: the International Skin Imaging Collaboration (ISIC), which comprises dermoscopic pictures necessary for the classification of benign lesions and melanoma, and MED-NODE, which includes non-dermoscopic images to facilitate practical diagnostic applications. The approach uses deep learning, namely convolutional neural networks (CNNs) like ResNet50 and InceptionV4, which mimic neural processing to identify patterns of skin cancer in photos, High-level pattern recognition specific to malignant lesions is made possible by these architectures' extraction of multi-level features. Important methods include using labeled photos to train models for supervised learning and then comparing the model's performance to dermatologists. According to the results, CNNs can distinguish melanoma with up to 95% sensitivity and 76.7% specificity, often outperforming dermatologists in diagnosis. CNN-based machine learning models show potential as effective, non-invasive diagnostic tools that supplement dermatologists' expertise and improve skin cancer care outcomes by promoting early detection and lowering reliance on biopsies. Important methods include using labeled photos to train models for supervised learning and then comparing the model's performance to dermatologists. According to the results, CNNs can distinguish melanoma with up to 95% sensitivity and 76.7% specificity, often outperforming dermatologists in diagnosis accuracy. CNN-based machine learning models show potential as effective, non-invasive diagnostic tools that supplement dermatologists' expertise and improve skin cancer care outcomes by promoting early detection and lowering reliance on biopsies. Ding et al. (2024)[5] present an automated melanoma diagnosis technique. Using 68 pairs of dermoscopic images from the American Cancer Society (ACS) collection, following preprocessing (normalization and histogram equalization the Gray-Level Co-occurrence Matrix is used to extract texture-based information (GLCM). The optimal feature selection is obtained using a modified Archimedes Optimization Algorithm (DAOA), which combines dynamic transfer operators, opposition-based learning, and chaotic mapping to improve feature selection and classifier parameters. A Support Vector Machine (SVM) is used to classify the images as either benign or malignant tumors, then DAOA is used to further improve the classification. The model outperformed comparable techniques and showed promise for precise early melanoma diagnosis with 88% accuracy, 96% sensitivity, 81% specificity, 97% precision, and a 97% F- measure. Priyadharshini et al. [6] proposed the paper using a dataset of 300 photos from Kaggle and DermIS that have been divided into benign and malignant classes, the paper presents a hybrid ELM-TLBO model for melanoma identification. A median filter is used to reduce noise, for segmentation, fuzzy C-means is utilized, and PCA is used to extract features. The ELM-TLBO model outperformed models like SVM and DCNN with an accuracy of 93.18% by fusing TLBO optimization with ELM's rapid neural network. This model shows notable improvement with 89.72% precision, 92.45% recall, and 91.64% F1-score; further research in scalability and cloud-based systems is recommended. The current paper by Sing et. al[7], introduces a sophisticated melanoma detection model that is trained, tested, and validated using three datasets: PH2, ISIC 2017, and ISIC 2018. Among the preprocessing techniques are histogram equalization for contrast enhancement and the Dull Razor algorithm for hair removal, using a two-phase segmentation procedure—dynamic thresholding in the first phase and L-Function fuzzy logic in the second—allows for precise lesion segmentation.

Using a modified YOLO model with extra layers to capture local and global context, classification achieves over 95% accuracy, sensitivity, and specificity. In dermatology, the combination of preprocessing, fuzzy logic-based segmentation, and YOLO classification exhibits great promise for therapeutic use. Pamel et. al(2024)[8] study in This paper used the trained CNN models for the diagnosis of skin cancer using datasets such as HAM10000, ISIC, and PH2. With some models reaching up to 99% accuracy, CNN architectures like ResNet and DenseNet—which are optimized with algorithms like ADAM and SGD—are renowned for their exceptional accuracy. The three most important metrics—specificity, recall, and accuracy are essential because they guarantee accurate identification while reducing errors. However, dataset limitations, particularly skin tone variability, provide a barrier to the model's generalizability. According to the paper's conclusion, improving hybrid-trained models and resolving these dataset biases are crucial next stages for successful clinical application. Akanksha Maurya et. al(2024)[9] study in this paper combines Topological Data Analysis (TDA) using Persistence Homology and Deep Learning (DL) with a fine-tuned EfficientNet-B5 model for Basal Cell Carcinoma (BCC) diagnosis. The methodology integrates U-Net for segmenting telangiectasia, TDA features, and DL features via Random Forest ensemble learning. The hybrid model achieved an accuracy of 97.4% and an AUC of 0.995, demonstrating the effectiveness of incorporating telangiectasia features. Future work aims to include more clinical biomarkers and validate the approach with statistical significance testing. Parasca et. al[10] presented in this paper A non-invasive technique for defining tumor margins in squamous cell carcinoma (SCC) and basal cell carcinoma (BCC). The suggested method divides and categorizes tumor and healthy skin areas according to spectral characteristics by combining hyperspectral imaging (HSI) with a machine learning classifier called Spectral Angle Mapper (SAM). By employing region-based segmentation, the technique lowers computational complexity in contrast to pixel-level methods. With median AUC values of 0.8914 for SCC and 0.8930 for BCC, the study's evaluation of 11 patients revealed excellent performance and good tumor margin detection accuracy. This strategy is a useful preoperative tool because of its efficiency and objectivity. A tiny sample size, a narrow range of tumor forms, and no comparison with gold-standard histology are some of the shortcomings, though. The upcoming goal of future research is to increase the method's reliability for wider clinical applications and test it on bigger datasets. Tim Hartmann et al. (2024)[11] propose the use of a dataset of 130,000 clinical photos with biopsy confirmation to investigate the use of AI in melanoma detection. It analyzes lesion properties including asymmetry, color, and boundaries using convolutional neural networks (CNNs) and designs like ResNet. Although the models' sensitivity was above 90% and their accuracy was dermatologist-level, issues like overdiagnosis and data generalization still exist. The paper emphasizes how future developments could lead to better dataset variety and human-AI cooperation. Bello et.al(2024)[12] purposed in this paper categorization of skin cancer by transfer learning methods. Using preprocessing and data augmentation, the fine-tuned DenseNet-121 model outperformed ResNet-34, VGG-16, and EfficientNetB0 in terms of accuracy (87%). The work emphasizes the need for additional varied datasets to enhance generalizability and clinical acceptance, despite its computational efficiency and strong performance. Ashutosh Lembhe et al. (2022)[13] offer a deep learning strategy to enhance their study's skin cancer diagnosis. Improvement in Skin Cancer Detection with Convolutional Neural Networks and Image Super Resolution, before feeding the images into CNN models (VGG16, ResNet, and InceptionV3) for classification, the authors improve image quality using Image Super Resolution (ISR) with a GAN framework.

They do this by splitting the ISIC dataset of skin lesion images, for teaching and testing purposes, which has 1800 benign and 1497 malignant melanoma photos in an 80:20 ratio. This approach significantly improved the models' accuracy: VGG16 increased from 54.55% to 70.17%, ResNet increased from 72.72% to 86.57%, and InceptionV3 increased from 83.48% to 91.26%. The findings demonstrate ISR’s capacity to increase diagnostic precision and imply that it may find broad use in other areas of medical imaging. Future research attempts to integrate more extensive clinical data and improve accuracy through sophisticated segmentation. Ravi Manne et al.(2020)[14] findings demonstrate ISR’s capacity to increase diagnostic precision and imply that it may find broad use in other areas of medical imaging. Future research attempts to integrate more extensive clinical data and improve accuracy through sophisticated segmentation. This work, Opportunities, and Vulnerabilities, examines different CNN architectures for the classification of skin cancer, with particular attention to datasets such as HAM10000, DermNet, and ISIC. Pretrained CNNs like DenseNet201 and ResNet152 offer remarkable accuracy, sometimes outperforming dermatologists, according to key findings. DenseNet201 even achieved an AUC of 99.30% for basal cell cancer. The authors draw attention to the model’s susceptibility to color changes, rotation, and hostile attacks, but they also stress the need for highlight in this review, which also calls for improvements in model reliability. Rajendran et. al[15] presented in this paper ISIC2017 and HAM10000 databases, which contain dermoscopic pictures of skin lesions for melanoma and other disorders, are used in this work. The process includes a Gated Recurrent Unit (GRU) for classification, U-Net for segmentation, MobileNet for feature extraction, and Bilateral Filtering for noise reduction. GRU hyperparameters are optimized using Cat Swarm Optimization (CSO). These elements are included in the architecture for effective and precise skin cancer diagnosis. With an accuracy of 97.44% on ISIC2017 and 98.48% on HAM10000, the model outperforms current methods in terms of sensitivity, specificity, and F1 score. Naqvi et. al[16] presented in this paper The paper emphasize how different datasets, including PH2, ISIC 2016–2020, and HAM10000, are used for skin cancer diagnosis using preprocessing, segmentation, feature extraction, and classification techniques. DenseNet, ResNet, and NASNet are common architectures that have shown excellent performance on particular datasets. For instance, NASNet obtained 97.7% accuracy on ISIC 2020, whereas DenseNet obtained 98.33% accuracy on HAM10000 with an F1 score of 0.96. Lightweight models like Squeeze-MNet scored an outstanding 99.36% accuracy on ISIC datasets, demonstrating potential for deployment on IoT devices. Across studies, classification accuracy was greatly increased by methods like ensemble frameworks, data augmentation, and transfer learning.

Table 1 presents a comparative analysis of various deep learning architectures, datasets, and techniques applied in recent studies, highlighting model choices and accuracy outcomes. It also outlines each study's future research directions and contributions by different author

Table 1. Literature Review Table with lots of information Findings from performance evaluations of cutting-edge algorithms for skin lesion segmentation

Author Name	Year	Techniques Used	Advantages	Drawbacks	Computational Costs	Datasets Used	Results
Gosh et.al [1]	2024	Hybrid CNN (VGG16 + ResNet50), data augmentation, class weights	High training (98.75%) and validation (97.50%) accuracy; effective for handling class imbalance	Limited testing on real-world clinical datasets	Moderate, depending on the hybrid model	Dataset with 3000 images (9 skin conditions)	Precision: 97.60%, Recall: 97.55%, F1 Score: 97.58%
Venugopal et al. [2]	2023	Modified EfficientNetV2-M, transfer learning, data augmentation	High accuracy (97.62%) for multiclass classification; robust with large datasets	Requires further GAN-based synthetic data for balance	High Computational cost for fine-tuning EfficientNet	ISIC 2018, 2019, 2020 datasets	Binary Classification: 99.23% accuracy; Multiclass Classification: 97.62%
Gouda et al. [3]	2022	CNN with Enhanced Super-Resolution GAN (ESRGAN), transfer learning	Improved image quality leads to better feature extraction; robust preprocessing techniques	Lower accuracy (85.7%) compared to state-of-the-art methods	Moderate	ISIC 2018 dataset	Accuracy: 85.7% (InceptionV3 model), Precision comparable to dermatologist-level competence
Kinnor das et.al [4]	2021	Machine learning, Deep Learning, CNNs (ResNet50, InceptionV4, Transfer Learning, ESRGAN for augmentation	Non-invasive diagnosis; enhances dermatologists' accuracy; enables early detection; reduces reliance on biopsies	A specificity of 76.7% is lower compared to sensitivity, indicating possible false positives	Moderate	ISIS (dermoscopic images), MED-NODE (non-dermoscopic images)	CNN achieved up to 95 %, 76.7% specificity; 85.7% accuracy in distinguishing benign and malignant lesions
Ding et al. [5]	2024	Archimedes Optimization Algorithm (DAOA), SVM classification	High sensitivity (96%) and precision (97%); texture-based feature extraction effective	Specificity is comparatively low (81%)	Moderate	ACS dataset (68 image pairs)	Accuracy: 88%, Sensitivity: 96%, Specificity: 81%, Precision: 97%, F1 Score: 97%
Priyadharshini et al. [6]	2023	Hybrid ELM-TLBO model, PCA for feature extraction, Fuzzy C-Means clustering	Outperformed traditional models; lightweight design for scalability	Limited dataset; accuracy drops with high noise levels	Low computational cost	Kaggle, DermIS datasets (300 images)	Accuracy: 93.18%, Precision: 89.72%, Recall: 92.45%, F1 Score: 91.64%
Singh et al. [7]	2023	Modified YOLO, histogram equalization, L-R fuzzy logic segmentation	High sensitivity and specificity (95%+); suitable for real-time applications	Requires optimization for mobile deployment	High computational cost	ISIC 2017, ISIC 2018, PH2 datasets	Sensitivity and Specificity: 95%+, High IoU values
Hermosila et.al [8]	2024	Convolutional Neural Networks (CNNs), Including ResNet, DenseNet, and Hybrid approaches with transfer learning. Techniques: GANs for synthetic data, Adam/SGD/RSProp optimization,pre-processing (noise reduction, normalization, edge detection).	High accuracy 99.9% depending on the architecture and dataset.	Requires further GAN-based synthetic data for balance	Moderate	HAM10000, ISIC, PH2 datasets; used for training and Validation of Diagnostic models.	Accuracy Ranges from 88% to 99.9% depending on architecture and dataset.

Akanksha Maurya et. al [9]	2024	Combines Topological Data Analysis (TDA) using Persistence Homology and fine-tuned EfficientNet-B5. U-Net for telangiectasia segmentation; EfficientNet-B5 for feature extraction; Ensemble learning with Random Forest.	Enhanced accuracy with telangiectasia features, computationally efficient.	Limited dataset diversity; vessel masks annotated by a single dermatologist	The TDA feature computation is resource-efficient and does not require high-performance GPUs, reducing overall computational cost.	2000 images from HAM10000, ISIC 2019, and NIH datasets.	Accuracy: 97.4%, AUC: 0.995 for BCC diagnosis.
Sorin Viorel Parasca et.al [10]	2024	Hyperspectral Imaging (HSI) combined with Spectral Angle Mapper (SAM) classification for tumor margins	Non-invasive, objective, and efficient in preoperative assessment of tumor margins	Small sample size, limited to specific tumor types, and lack of comparison with gold-standard histology	Reduced by segment-based classification instead of pixel-level analysis	patients (6 with Basal Cell Carcinoma (BCC) and 5 with Squamous Cell Carcinoma (SCC))	Median AUC: 0.8914 for SCC, 0.8930 for BCC; high accuracy in tumor margin delineation
Tim Hartmann et.al [11]	2024	Artificial Intelligence, Convolutional Neural Networks (CNNs), supervised learning, anomaly detection, dimensionality reduction. Residual Networks (ResNet), Autoencoders	High accuracy comparable to dermatologists, potential for early detection, reduced healthcare costs, and applications in teledermatology and rural healthcare.	Limited dataset diversity, risks of overdiagnosis, lack of generalizability, challenges with real-world image quality, and reliance on optimized datasets	Requires high computational power for training large datasets and deep network architectures.	130,000 labeled clinical images representing over 2,000 skin conditions, all biopsy-proven for reliability.	Demonstrated effectiveness in classifying malignant melanoma and benign lesions under optimized conditions. Sensitivity above 90% for certain models like DermAI, with dermatologist-level performance in some cases.
Abayomi Bello et.al [12]	2024	Transfer learning using DenseNet-121, ResNet-34, VGG-16, Inception v3, EfficientNetB0, and traditional methods like CNN, SVM, and Random Forest.	Improved accuracy and feature reuse using DenseNet-121; effective data augmentation enhanced robustness and reduced overfitting.	Limited dataset diversity (did not cover all skin colors); potential bias due to dataset imbalance; constrained generalizability to all real-world scenarios.	Moderate due to fine-tuning and additional dense layers; models required 20 epochs with data augmentation and optimized training parameters.	Combined HAM10000 and ISIC 2020 datasets, totaling 3,297 images (1,800 benign, 1,497 malignant), with 5,636 images post-augmentation.	
Ashutosh Lembhe et.al [13]	2023	Image Super Resolution (ISR) using GANs to enhance low-resolution images combined with CNN-based models (VGG16, ResNet, Inception V3) for classification.	Increased classification accuracy through ISR, enhanced image quality for better feature extraction, and potential to reduce storage costs by enabling efficient resolution enhancement.	ISR introduces additional computational steps; dataset diversity limited to ISIC archives; performance limited by hardware and dataset size.	Moderate computational requirements, including Windows 10 PC, Core i5 (9th Gen), 16 GB RAM, and GTX 1650 GPU. Training involved 100 epochs with a batch size of 32.	ISIC dataset with 3,297 images (1,800 benign, 1,497 malignant) used with 80:20 train-test split; additional data augmentation applied.	The best accuracy achieved was 91.26% using Inception V3 with ISR, showing significant improvement over models without ISR (e.g., VGG16 improved from 54.55% to 70.17%).

Ravi Manne et.al [14]	2020	Deep learning with CNNs, transfer learning, hybrid models (e.g., CNN with Fisher Vector Encoding, SVM, KNN), and end-to-end learning.	Higher accuracy than dermatologists in many cases; CNNs reduced the need for manual feature extraction; effective for both benign and malignant classification.	Vulnerabilities to adversarial attacks, image perturbations, and dataset bias (e.g., limited representation of dark-skinned individuals); varying performance based on image acquisition methods.	Varies by approach; some studies used resource-intensive models like DenseNet201 and ResNet152. Techniques like transfer learning reduced computational demands by leveraging pretrained weights.	Multiple datasets reviewed, including HAM10000, ISIC 2018, and DermoFit. Dataset sizes varied, with some studies using up to 19,398 images across different lesion types.	Accuracy varied across methods: DenseNet201 achieved 98.16% for HAM10000, ResNet models reached 96% for diverse datasets, and CNN-based hybrid methods improved lesion detection sensitivity and specificity.
Vijay Arumugam Rajendran et.al [15]	2024	Bilateral Filtering for noise reduction, U-Net for segmentation, MobileNet for feature extraction, GRU for classification, and Cat Swarm Optimization (CSO) for hyperparameter tuning	High accuracy, reduced noise, effective segmentation, lightweight model, and optimized hyperparameters	Computational complexity due to integrated algorithms and challenges in real-time processing and generalization	MobileNet reduces computational cost through lightweight architecture; GRU with CSO optimizes efficiency	ISIC2017 and HAM10000 datasets (2000 samples, 3 classes)	Achieved 97.44% accuracy on ISIC2017 and 98.48% accuracy on HAM10000 dataset
Naqvi et al [16]	2023	Various CNN models (e.g., ResNet, VGG, Xception)	High accuracy for skin lesion detection	Limited datasets with skin color diversity	Training on GPU clusters or high-end CPUs	HAM10000, ISIC datasets	Accuracy up to 98.70% (HAM10000)

Kumar Lilhore et. al[17] proposed the paper HAM-10000 datasets used in the study to suggest a hybrid model for skin cancer detection. It uses an enhanced MobileNet-V3 for feature extraction and U-Net for segmentation, with hyper parameters adjusted using Bayesian techniques. Data augmentation techniques like flipping, rotation, and brightness adjustment are part of pre-processing. The model outperformed current techniques in the diagnosis of skin cancer, with 98.86% accuracy, 97.84% precision, and 96.35% sensitivity. Midasala et. all(2024)[18] propose the work presents MFEUsLNet, an AI model for skin cancer diagnosis, using the ISIC-2020 dataset of 33,126 dermoscopic pictures of seven distinct types of skin cancer, including melanoma and basal cell carcinoma. Important texture and color information is extracted using the Gray-Level Co-occurrence Matrix (GLCM) and Redundant Discrete Wavelet Transform (RDWT). Lesions are segmented using K-means clustering (USL-KMC), and a bilateral filter is applied to reduce noise. A Recurrent Neural Network (RNN) performs the classification, outperforming similar models with 99.18% accuracy, 99.82% precision, 99.91% recall, and a 99.75% F1-Score. This system expertly combines RNN-based classification with preprocessing, segmentation, and multilevel feature extraction to deliver a dependable automated skin cancer diagnosis technique. Alqarafi et. al[19] proposed The study uses the Multi-scale GC-T2 framework to detect skin cancer using the DermIS and DermQuest databases. A tri-level feature fusion module, a tri-movement attention mechanism for dimensionality reduction, and a Multi-scale Graph Convolutional Network (M-GCN) for feature extraction are important parts. In preprocessing, the Median Enhanced Weiner Filter is used to reduce noise, while the Enriched Manta-Ray Optimization Algorithm is used to improve quality. Deep Reinforcement Learning (AdDNet) and HAar-U-Net are combined in semantic segmentation. With 98.9% accuracy on DermIS and 97.93% accuracy on DermQuest, the suggested model demonstrates its promise. The system will be improved in future work and used for various medical imaging applications and bigger datasets. Kandhro et. al[20] proposed the paper ISIC 2020 dataset, which includes more than 33,000 photos of skin lesions, is used in this study. In order to diagnose skin cancer more accurately, the researchers added layers and features to the VGG19[21] model and contrasted it with other models such as ResNet and DenseNet[22]. The suggested E-VGG19 model performed better in identifying benign and malignant tumor techniques[23], with an accuracy of about 88%. Real-time detection, sophisticated segmentation, and maybe the application of quantum computing to improve model efficacy in skin cancer diagnosis are some potential future possibilities.

After reviewing these papers, I found some limitations. There are significant drawbacks to deep learning models for skin cancer diagnosis. Inadequate dataset size and diversity frequently result in biases and make it more difficult to generalize across various populations and lesion types. Many models are unsuitable for real-time or resource-constrained clinical contexts due to their high computing costs. Variations in sensitivity and specificity can affect reliability by producing false positives or negatives. Since many models are trained in controlled settings, they are less applicable in situations where image quality varies in the real world. With smaller datasets, overfitting is still an issue, and model validation is made more difficult by the lack of standardization in preprocessing and evaluation. The necessity for scalable, dependable, and therapeutically appropriate solutions is highlighted by these difficulties.

The evaluated studies demonstrate notable progress in the application of deep-learning models for the detection of skin cancer, with increased preprocessing techniques, sophisticated feature selection algorithms, and hybrid architectures leading to gains in accuracy, sensitivity, and precision. However, there are still issues because there hasn't been enough training on a variety of datasets, which limits how broadly the models can be applied. It is uncertain how to apply computationally demanding models, such as EfficientNet, in resource-constrained environments, despite their superior performance. Furthermore, even if outcomes have improved when techniques like fuzzy logic and transfer learning are combined, problems like high processing costs and decreased specificity still exist. Several drawbacks are shown by reviewing the literature: models are less appropriate to diverse populations and real-world contexts due to biases caused by inadequate dataset variety. With limited datasets, many models are prone to overfitting and are computationally costly. Reliability is also impacted by variations in sensitivity and specificity. These difficulties highlight the requirement for clinically viable, scalable, and dependable models that take into account a variety of skin types and lesion complexity.

3. Common dataset of skin lesion photos for model assessment

3.1 PH2

Used extensively for lesion segmentation and classification tasks, this collection of dermoscopic pictures includes thorough annotations and diagnosis labels (common nevus, atypical nevus, melanoma) Dermoscopic images having a 768x560 pixel resolution can be found in PH2[24]. It includes 200 melanocytic lesion dermoscopic images, Including 40 melanomas, 80 atypical nevi, and 80 common nevi. https://www.kaggle.com/datasets/athina123/ph2dataset/data

3.2: ISIC (International Skin Imaging Collaboration) Archive

ISIC[25] is the largest publicly accessible dataset, comprising more than 25,000 photos of skin lesions, such as nevi, keratoses, and melanomas, Together with expert comments and related metadata.
https://gallery.isic-archive.com/#!/topWithHeader/onlyHeaderTop/gallery?filter=%5B%22diagnosis%7Csquamous%20cell%20carcinoma%22%5D

3.3 HAM10000 (Human Against Machine with 10000 Cases)

Contains 10,015 dermatologist-annotated dermoscopic photos of seven distinct skin lesion types that are widely used in deep learning and classification studies. https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000/discussion/173439#963721

4. Methodology Review for advancing the Field

Concrete methodologys for furthering the field, leveraging transfer learning to expand the model to other medical imaging applications, such as CT and mammography, enhancing robustness with other datasets and GANs, and verifying the model in actual clinical settings through partnerships with healthcare organizations. It also recommends using explainable AI methods like Grad-CAM to guarantee transparency and promote clinician trust, as well as integrating multi-modal data, such patient metadata, to improve prediction capability. The goal of these changes is to offer doable actions for significant domain contributions.

5. Extended Conversation on Associated Research: Placing the Present Study in the Larger Scientific Framework

Recent developments in deep learning and machine learning have drastically changed the field of medical image analysis, especially in the area of skin cancer detection. Conventional diagnostic methods, which frequently depend on clinical knowledge and dermoscopic analysis, are sensitive to subjectivity and variability. Early research merged hand-crafted features like color, texture, and form descriptors with traditional machine-learning approaches like Support Vector Machines (SVM) and k-Nearest Neighbors (k-NN). Although these techniques showed some degree of effectiveness, their reliance on feature engineering limited their capacity to generalize across various lesion types. By automating feature extraction and achieving exceptional accuracy in tasks like lesion recognition and segmentation, Convolutional Neural Networks (CNNs) transformed image analysis. Deeper and more complex models, such as ResNet, DenseNet, and Inception, were made possible by architectures like AlexNet and VGG. These architectures improve their ability to mimic the intricate patterns found in dermoscopic images by utilizing deeper layers, residual connections, and multi-scale feature extraction. The benefits of compound scaling in striking a balance between model accuracy and computational efficiency have been further illustrated by recent research using EfficientNet and its variations.

To capitalize on the capabilities of various networks, a number of hybrid models that combine multiple CNN designs have also been developed. To enhance feature variety and classification performance, for example, research has looked into combining VGG16 and ResNet50. Since there are frequently few labeled datasets in the field of medical imaging, methods like transfer learning have proven especially helpful. Rotation, scaling, and flipping are examples of data augmentation techniques that are frequently used to improve model generalization. Even with these developments, problems still exist. Given that malignant lesions are usually underrepresented in datasets, class imbalance is still a serious problem. Due to their propensity to favor benign classes, models trained on unbalanced datasets have a lower sensitivity for melanoma detection. Furthermore, model generalization is hampered by the variability in dermoscopic image collections, which results from variations in resolution, illumination, and lesion characteristics. By suggesting a hybrid CNN model that combines VGG16 and ResNet50, utilizing their complementary strengths in feature extraction and classification, the current study expands on these developments. The objective of this study is to enhance the accuracy and robustness of skin lesion classification by tackling important issues such as class imbalance using class weights and sophisticated preprocessing. Additionally, testing on a variety of datasets, such as ISIC and HAM10000, guarantees that the model is evaluated on a wide range of intricate real-world instances, establishing it as a noteworthy advancement in the field of computer-aided skin cancer diagnosis.

6. Result analysis and discussion:

It is clear from the table's study that deep learning models routinely perform better than conventional machine learning methods in the segmentation of skin lesions and the diagnosis of skin cancer. The best accuracy (97.50%) was attained by hybrid[26] and ensemble models, such as the VGG16 and ResNet50 combination, which also showed outstanding performance on precision, F1 score, and recall measures. Better testing accuracy was also demonstrated by models such as EfficientNetV2-M[27], especially on benchmark datasets like ISIC2019 (97.62%). Model performance was also greatly enhanced by the use of Image Super Resolution (ISR)[28]; for example, Inception V3 achieved a noteworthy accuracy of 91.26% when augmented with ISR. The improvements brought about by deep learning architectures were demonstrated by the much poorer performance of more conventional techniques, such as SVM and Logistic Regression. These results highlight the significance of sophisticated.

Table 2 evaluates a variety of machine learning and deep learning models for detecting or segmenting skin lesions, with their performance evaluated in terms of accuracy during training and testing. It also highlights improvements using techniques like Image Super Resolution (ISR) and hybrid approaches. Models like EfficientNet and hybrid architectures (e.g., VGG16 + ResNet50) demonstrate high accuracy in detecting skin cancer.

Table 2. Findings from performance evaluations of cutting-edge algorithms for skin lesion segmentation

Paper	Model	Accuracy	
		Training	Testing
A Study on the Application of Machine Learning and Deep Learning Techniques for Skin Cancer Detection [1]	DenseNet121	99.51%	91.82%
	VGG16	-	70.03%
	Linear SVM	-	74.33%
	Nearest Neighbor (KNN)	-	81.56%
	Decision Tree	-	68.67%
	ResNet50	88.89%	88.89%
	Hybrid DL Model (VGG16 & ResNet50)	98.75%	97.50%
	precision score	-	97.60%
	F1 Score	-	97.58%
	Recall	-	97.55%
A deep neural network using modified EfficientNet for skin cancer detection in dermoscopic images [2]	EfficientNet-B4	-	96.95%(ISIC2020)
	EfficientNetV2-M	-	97.62%(ISIC2019)
A novel hybrid Extreme Learning Machine and Teaching–Learning Based Optimization algorithm for skin cancer detection [6]	Support Vector Machine	-	78.11%
	Deep Convolutional Neural Network	-	83.25%
	Logistic Regression	-	78.21%
	ELM-TLBO	-	93.18%
Enhancement in Skin Cancer Detection using Image Super Resolution and Convolutional Neural Network [9]	VGG16(Without ISR)	-	54.55%
	ResNet (Without ISR)	-	72.72%
	Inception V3(Without ISR)	-	83.48%
	VGG16(With ISR)	-	70.17%
	ResNet (With ISR)	-	86.57%
	Inception V3(With ISR)	-	91.26%
Detection of Skin Cancer Based on Skin Lesion Images Using Deep Learning [3]	CNN	-	83.18%
	Resnet50	-	83.64%
	InceptionV3	-	85.76%
	Inception ResNet	-	84.09%

7. Conclusion of Future Directions:

The paper emphasizes how deep learning models, especially hybrid CNN constructions like VGG16 and ResNet50, are promising to improve the early detection of skin cancer. Research indicates that these models can outperform conventional diagnostic techniques due to their high accuracy, precision, and recall. This study highlights the importance of diverse and well-annotated dermoscopic datasets in creating efficient models by looking at several datasets, including the ISIC Archive, HAM10000[29], and PH2. Combining these cutting-edge models with computer-aided diagnostic (CAD)[30] systems in clinical settings may result in earlier diagnosis, improved patient outcomes, and lower mortality rates related to melanoma and other skin cancers, as evidenced by the studies' continuous success. The review also notes important issues, such as the requirement that models generalize to various clinical contexts and demographics. Future studies should focus on enhancing model robustness, investigating a variety of training datasets, and creating frameworks for implementation in practical settings to overcome these problems. Overall, this review encourages further innovation to guarantee efficiency and accessibility Using deep learning-based CAD systems in practical uses, supporting their continued development as revolutionary tools in dermatology.

Although there are many chances for development and expansion, this study's hybrid deep-learning model has shown great promise for identifying skin cancer. Developing a generally applicable and therapeutically useful diagnostic tool requires extending the research to more medical imaging applications, enhancing model robustness using a variety of datasets, and conducting real-world validation. Future studies can push the limits of AI in medical diagnostics by tackling these issues, which will eventually improve healthcare outcomes and provide more individualized treatment alternatives.

References

[1] Ghosh, Hritwik, Irfan Sadiq Rahat, Sachi Nandan Mohanty, J. V. R. Ravindra, and Abdus Sobur. "A Study on the Application of Machine Learning and Deep Learning Techniques for Skin Cancer Detection." International Journal of Computer and Systems Engineering 18, no. 1 (2024): 51-59.

[2] Venugopal, Vipin, Navin Infant Raj, Malaya Kumar Nath, and Norton Stephen. "A deep neural network using modified EfficientNet for skin cancer detection in dermoscopic images." Decision Analytics Journal 8 (2023): 100278.]

[3] Gouda, Walaa, et al. "Detection of skin cancer based on skin lesion images using deep learning." Healthcare. Vol. 10. No. 7. MDPI, 2022.

[4] Das, Kinnor, et al. "Machine learning and its application in skin cancer." International Journal of Environmental Research and Public Health 18.24 (2021): 13409.

[5] Ding, Huan, Qirui Huang, and Ahmed Alkhayyat. "A computer-aided system for skin cancer detection based on a Developed version of the Archimedes Optimization algorithm." Biomedical Signal Processing and Control 90 (2024): 105870.

[6] Priyadharshini, N., Selvanathan, N., Hemalatha, B., and Sureshku- mar, C. 2023. "A Novel Hybrid Extreme Learning Machine and Teaching–Learning-Based Optimization Algorithm for Skin Cancer Detection." Healthcare Analytics 3: 100161.

[7] Singh, Sumit Kumar, Vahid Abolghasemi, and Mohammad Hossein Anisi. "Fuzzy logic with deep learning for detection of skin cancer." Applied Sciences 13.15 (2023): 8927.

[8] Hermosilla, Pamela, et al. "Skin Cancer Detection and Classification Using Neural Network Algorithms: A Systematic Review." Diagnostics 14.4 (2024): 454.

[9] Maurya, Akanksha, et al. "Hybrid Topological Data Analysis and Deep Learning for Basal Cell Carcinoma Diagnosis." Journal of Imaging Informatics in Medicine 37.1 (2024): 92-106.

[10] Parasca, Sorin Viorel, et al. "Hyperspectral imaging with machine learning for in vivo skin carcinoma margin assessment: a preliminary study." Physical and Engineering Sciences in Medicine (2024): 1-12.

[11] Hartmann, Tim, et al. "Basic principles of artificial intelligence in dermatology explained using melanoma." JDDG: Journal der Deutschen Dermatologischen Gesellschaft 22.3 (2024): 339-347.
[12] Bello, Abayomi, Sin-Chun Ng, and Man-Fai Leung. "Skin Cancer Classification Using Fine-Tuned Transfer Learning of DENSENET-121." Applied Sciences 14.17 (2024): 7707.
[13] Lembhe, Ashutosh, et al. "Enhancement in Skin Cancer Detection using Image Super Resolution and Convolutional Neural Network." Procedia Computer Science 218 (2023): 164-173.
[14] Manne, Ravi, Snigdha Kantheti, and Sneha Kantheti. "Classification of Skin cancer using deep learning, ConvolutionalNeural Networks-Opportunities and vulnerabilities-A systematic Review." International Journal for Modern Trends in Science and Technology, ISSN 6 (2020): 2455-3778.
[15] Rajendran, Vijay Arumugam, and Saravanan Shanmugam. "Automated Skin Cancer Detection and Classification using Cat Swarm Optimization with a Deep Learning Model." Engineering, Technology & Applied Science Research 14.1 (2024): 12734-12739.
[16] Naqvi, Maryam, et al. "Skin cancer detection using deep learning—a review." Diagnostics 13.11 (2023): 1911.
[17] Kumar Lilhore, Umesh, et al. "A precise model for skin cancer diagnosis using hybrid U-Net and improved MobileNet-V3 with hyperparameters optimization." Scientific Reports 14.1 (2024): 4299.
[18] Midasala, Vasuja Devi, et al. "MFEUsLNet: Skin cancer detection and classification using integrated AI with multilevel feature extraction-based unsupervised learning." Engineering Science and Technology, an International Journal 51 (2024): 101632.
[19] Alqarafi, Abdulrahman, et al. "Multi-scale GC-T2: Automated region of interest assisted skin cancer detection using multi-scale graph convolution and tri-movement based attention mechanism." Biomedical Signal Processing and Control 95 (2024): 106313.
[20] Kandhro, Irfan Ali, et al. "Performance evaluation of E-VGG19 model: Enhancing real-time skin cancer detection and classification." Heliyon 10.10 (2024).
[21] Faghihi, Amir, Mohammadreza Fathollahi, and Roozbeh Rajabi. "Diagnosis of skin cancer using VGG16 and VGG19 based transfer learning models." Multimedia Tools and Applications 83.19 (2024): 57495-57510.
[22] Bello, Abayomi, Sin-Chun Ng, and Man-Fai Leung. "Skin Cancer Classification Using Fine-Tuned Transfer Learning of DENSENET-121." Applied Sciences 14.17 (2024): 7707.
[23] R. Agrawal and M. Kumar, "Deep Learning Framework for Banana Shelf Life Classification based on Ripening: BSLC," 2023 6th International Conference on Information Systems and Computer Networks (ISCON), Mathura, India, 2023, pp. 1-5, doi: 10.1109/ISCON57294.2023.10112153.
[24] Latha, M., et al. "Enhancing Skin Cancer Classification on the PH2 Dataset Through Transfer Learning Technique." International Research Journal on Advanced Engineering Hub (IRJAEH) 2.03 (2024): 500-507.
[25] Imran, Azhar, et al. "Skin cancer detection using combined decision of deep learners." IEEE Access 10 (2022): 118198-118212.
[26] Daghrir, Jinen, et al. "Melanoma skin cancer detection using deep learning and classical machine learning techniques: A hybrid approach." 2020 5th international conference on advanced technologies for signal and image processing (ATSIP). IEEE, 2020.
[27] Islam, Niful, et al. "Leveraging Knowledge Distillation for Lightweight Skin Cancer Classification: Balancing Accuracy and Computational Efficiency." arXiv preprint arXiv:2406.17051 (2024).
[28] Adamu, Shamsuddeen, et al. "The future of skin cancer diagnosis: a comprehensive systematic literature review of machine learning and deep learning models." Cogent Engineering 11.1 (2024): 2395425.
[29] Batista, Lucas G., Pedro H. Bugatti, and Priscila TM Saito. "Classification of skin lesion through active learning strategies." Computer Methods and Programs in Biomedicine 226 (2022): 107122.
[30] Malibari, Areej A., et al. "Optimal deep neural network-driven computer aided diagnosis model for skin cancer." Computers and Electrical Engineering 103 (2022): 108318.