Deep Learning-Based Multi-Class Skin Cancer Detection: Methods, Challenges, and Future Directions

G. Nandhini*, T. Ananth Kumar, P. Kanimozhi
Department of Computer Science and Engineering IFET College of Engineering, Villupuram, India
*Corresponding Author Email: nandhinigopu55@gmail.com

Received: 10.04.2026 Revised: 14.05.2026 Accepted: 18.05.2026 Published: 20.05.2026

Abstract - Skin cancer is one of the most common cancers in the world and late diagnosis can result in sub-optimal treatment outcomes and mortality. The skin cancer incidence rate is steadily rising annually, which makes it even more crucial to develop automated diagnostic tools that can assist dermatologists in making correct and timely choices. In recent years, Artificial Intelligence (AI), including deep learning and machine learning have been increasingly adopted for the analysis of dermatoscopic skin images for skin cancer detection. This paper summarizes all the recently available techniques for multi-class skin cancer detection based on AI, including the datasets utilized by researchers, the deep learning architecture developed and the different classification methods proposed. The strengths and limitations of each method are also considered, and the major challenges and open research issues are explored. The primary aim of this work is to provide researchers with a clear depiction of the current state of the field and to highlight future research priorities.

Keywords: Skin Cancer Detection; Deep Learning; Multi-Class Classification; Convolutional Neural Networks, Explainable Artificial Intelligence

1.Introduction

The global burden of skin cancer continues to rise, with increasing numbers of cases and deaths reported across all continents. Among skin cancers, melanoma is the most aggressive and potentially fatal type, while nonmelanoma skin cancers such as basal cell and squamous cell carcinoma are more prevalent globally. In Europe, rates of melanoma are higher than anywhere else in the world. Australia and New Zealand also reported the highest melanoma rates worldwide, while Asia, despite having relatively lower incidence rates, records higher melanoma-related mortality compared to many other regions [1]. These regional differences demonstrate the continued increase in the worldwide burden of skin cancer and the need for better diagnostic methods for melanoma [2-4]. Early detection of skin cancer plays a critical role in improving patient survival rates and treatment outcomes. Dermatologists have traditionally relied on visual examination and dermoscopy to evaluate skin lesions by assessing asymmetry, border irregularity, color variation, and diameter. While dermoscopy allows clinicians to view features beneath the skin surface, the final diagnosis remains subjective and highly dependent on the clinician's experience and expertise. Since dermoscopic images are often interpreted differently by different clinicians, misdiagnosis remains a significant concern, particularly during the early stages of melanoma when visual cues are subtle and less distinctive. These limitations have driven growing interest in automated Computer-Aided Diagnosis (CAD) systems designed to support dermatologists in making more accurate and consistent diagnostic decisions [5]. Over the last few years, the introduction and advancement of AI have dramatically changed about how physicians utilize images for analysis and diagnosis. For example, deep learning algorithms have created very powerful deep learning models that can be used to analyze dermoscopic images by extracting complex patterns, thereby allowing for faster and more accurate diagnosis of skin cancer through automated methods. As such, when physician specialists are not available, the use of these types of systems can help provide a faster and more accurate diagnosis. These advantages make AI-based diagnostic systems particularly valuable in low-resource environments. In addition, the use of AI technology within the clinical workflow may provide better decision support and will enable more widespread screening using large populations [6 -10]. Building on these developments, recent advances in AI and machine learning (ML) models, specifically in the analysis of dermoscopic images, have greatly improved the classification of skin lesions [11,12].

Convolutional Neural Networks (CNNs) automatically learn to recognize an image's hierarchical features [13]. They also achieve higher accuracy than traditional handcrafted feature-based approaches in image classification. In many cases, using transfer learning in conjunction with traditional CNN methods, such as using VGG or DenseNet pre-trained architectures, further enhances the classification performance of the CNN; this is especially true for dermatology-related medical datasets that tend to be small [14]. There is a growing trend of evidence from systematic analysis that the diagnostic accuracy of deep learning models for the detection of melanoma is equivalent to that of dermatologists, and there is also considerable potential for the implementation of these models in clinical settings. In addition to their predictive capability, deep learning models need to be interpretable in order to ensure safe clinical use. Clinicians often raise concerns about deep learning models functioning as 'black boxes,' which undermines clinical trust in these systems. To counteract this challenge, a variety of Explainable Artificial Intelligence (XAI) techniques have been developed including Local Interpretable Model-agnostic Explanations (LIME) [15], Gradient-weighted Class Activation Mapping (Grad-CAM), Shapley Additive Explanations (SHAP), and Integrated Gradients. Recent studies explore hybrid architectures that merge Convolutional Neural Networks (CNNs) with Vision Transformers (ViTs), leveraging both local spatial features and global contextual representations to achieve improved multi-class classification performance, while preserving the explained transparency of both models' combined predictions [16]. This paper aims to provide a comprehensive analysis of recent advancements in deep learning, transfer learning, hybrid architectures, and explainable AI techniques for multi-class skin cancer detection. While researchers have made considerable progress in addressing these issues, developing a solution that generalizes across different populations and datasets, while maintaining computational efficiency and a balanced sensitivity-specificity trade-off, remains a significant challenge. Building reliable, interpretable, and clinically applicable systems for detecting multiple types of skin cancers continues to be an important research priority. This paper examines key findings and developments in this area, with particular focus on deep learning methods, transfer learning techniques, hybrid architectures, and explainable AI approaches [17-19]. Figure 1 illustrates the general pipeline of a deep learning-based skin cancer detection system, demonstrating the overall process of automated skin lesion classification which includes preprocessing, feature extraction, and final classification.

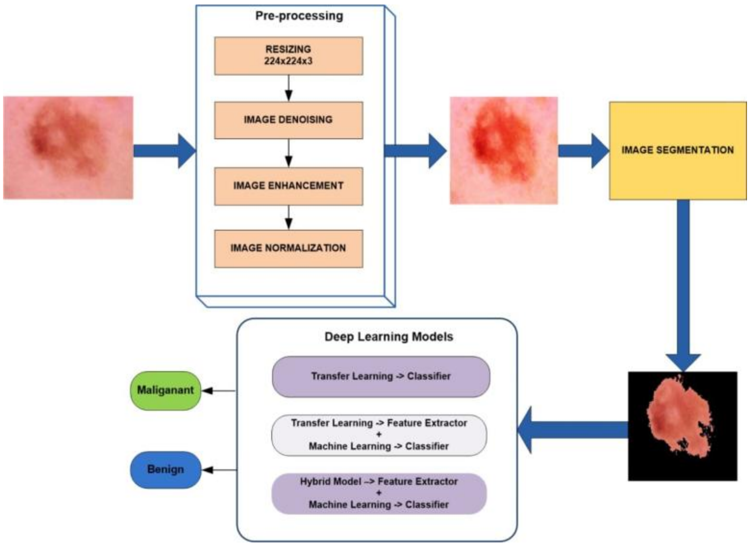

Figure 1 General pipeline of deep learning-based skin cancer detection

This paper focuses on several key aspects of deep learning-based skin cancer detection. It compiles the progression of automated detection methods from traditional image processing techniques to advanced deep learning architectures, including convolutional neural networks and attention-based models. Publicly available dermoscopic datasets are reviewed and assessed in terms of their strengths, weaknesses, and influence on model performance. The paper also examines the main challenges reported in current literature, such as class imbalance, generalization limitations, interpretability, and dataset bias. The clinical relevance of AI-based diagnostic systems and the barriers to their deployment in real-world healthcare settings are also discussed. Finally, existing gaps in the literature are identified and future research directions are proposed to improve the robustness, transparency, and fairness of multi-class skin cancer detection systems.

2. Methods used to detect Skin Cancer

Computational methods for automated detection of skin cancer have developed tremendously in the past two decades. In the past, handcrafted features were used along with classical machine learning classifiers but in recent years, deep learning architectures that learn features directly from dermoscopic images have gained considerable popularity. In this section, the key advances in this area are described, including traditional machine learning, convolutional neural networks, transfer learning, multi-class classification, attention mechanisms, vision transformers, hybrid architectures, multimodal fusion approaches, and explainable AI techniques.

2.1 Traditional Machine Learning Approaches

Hand engineered features derived from traditional machine learning approaches used to represent the skin and classify skin lesions as either benign or malignant were the primary area of research before the advent of deep learning convolutional networks. Few works has examined combining deep learning methods with handcrafted features (e.g., using RSurf descriptors or Local Binary Patterns (LBP)) to provide additional information about melanoma and provided improved accuracy of identifying melanoma than would be achievable using either type of features on their own [20]. P rior to the growth of deep learning, most researchers in the field relied on classical image processing methods such as Gaussian filtering for preprocessing, active contours for segmenting lesions, and Grey Level Co-occurrence Matrices (GLCMs) for texture features. After these objects were processed as above, they were typically classified using classical machine learning algorithms such as Support Vector Machines (SVM) and K-Nearest Neighbors (KNN). There was a substantial number of studies published that provided high levels of accuracy when using these classical machine learning approaches with respect to datasets like PH2 [21]. Numerous systematic reviews of computer-aided diagnosis have summarized both the transition from using classical machine learning approaches to using deep learning models [22]. In addition, they have noted major challenges associated with this transition, namely small-size datasets, selection bias associated with the images used in the datasets, and the ability to generalize across varying population types [23]. These earlier works have paved the way towards the emergence of new deep learning-based systems for the automated detection of skin cancer.

2.2 Convolutional Neural Networks (CNNs) for Skin Lesion Classification

Deep Learning has rapidly evolved over the last decade; therefore, CNNs are now being used to classify skin lesions automatically [24] instead of traditional machine learning methods [25]. The initial focus of these models was on binary classification tasks, with a primary focus on categorizing lesions as either malignant or benign [26,27]. These models displayed improved feature learning capabilities because they could automatically learn hierarchical representations using dermoscopy images; this allowed to overcome the challenges of using handcrafted feature extraction techniques [28,29]. In addition, a number of studies have developed enhanced CNN architectures that utilize new regularization techniques; therefore, these models can generalize better than previous models and were able to achieve high classification accuracy and a high degree of AUC-ROC across many types of comparisons between lesions. Comparisons were also made between CNN models and traditional classifiers (e.g., SVMs) with both types of approaches; CNN approaches tend to outperform traditional approaches with regard to both accuracy and robustness. Furthermore, systematic reviews of machine learning and deep learning methods for diagnosing melanoma found that there has been a surge in the number of architectures being used to develop CNNs (e.g., DenseNet and ResNet) and that these improvements provide new evaluations to be conducted on how the datasets are used and on how to validate the performance of CNN models [30]. The above studies reflect an important evolution to completely automated, deep learning-based skin cancer diagnostic systems [31,32]. However, these traditional approaches are limited by their dependence on manually crafted features and their inability to generalize across diverse datasets and populations. Figure 2 illustrates the basic architecture of a Convolutional Neural Network, demonstrating how input images pass through convolution and pooling layers for feature extraction before being classified in the fully connected layers.

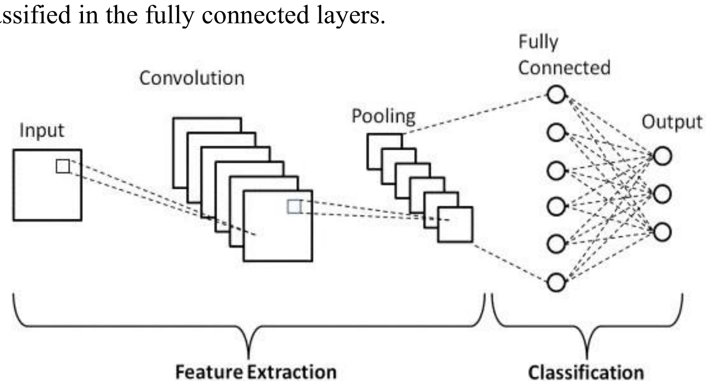

Figure 2 Basic architecture of a Convolutional Neural Network (CNN) in skin lesion analysis

2.3 Transfer Learning Techniques

Transfer learning is rapidly becoming a reliable method for assisting with skin lesion classification through the use of Deep Learning [33]. This approach has proven particularly useful in situations where there are no annotated datasets to confirm the diagnoses of skin lesions. Adapting pre-trained CNNs (convolutional neural networks) has been shown to significantly improve the ability of modern dermoscopy computers to detect skin lesions correctly [34]. Researchers have conducted several studies and have found that applying different pre-processing techniques like resizing and multi-scale cropping of images can lead to improved detection accuracy by retaining significant clinical details in cropped dermoscopy images [35]. The use of multi-scale and multi- network ensemble techniques, using different CNN training architectures, such as EfficientNet, SeResNeXt, DenseNet and ResNet, has resulted in higher overall accuracy on a variety of bench-mark datasets, including the International Skin Imaging Collaboration (ISIC) dataset [36]. Fine tuning these pre-trained models has been shown to produce even greater results, due to the introduction of additional dense layers and optimization of training parameters, with increased feature representation and model stability. Studies comparing the use of optimizers and batch sizes have also shown that well-tuned transfer learning CNN models provide excellent sensitivity and generalization performance [37]. Therefore, transferring knowledge from pre-trained models through transfer learning represents a viable and highly efficient way to create reliable automated skin cancer detection systems [38-40]. Despite their strong performance, CNN-based models often struggle with generalization when applied to datasets collected under different clinical conditions.

2.4 Multi-Class Classification Systems

As research progressed beyond binary classification, greater emphasis was placed on developing robust multi-class skin lesion classification systems capable of distinguishing among multiple disease categories simultaneously. Recent studies have introduced multi-model deep learning architectures that classify skin lesions into several clinically relevant categories using large-scale dermoscopic datasets. For instance, multi-model frameworks based on architectures such as Xception have demonstrated higher accuracy and strong Area Under the Receiver Operating Characteristic curve (AUROC) values when trained on tens of thousands of images, with additional transfer learning strategies improving class-specific performance across various skin conditions [41]. Systematic reviews of melanoma diagnosis further highlight that advanced deep learning architectures, including DenseNet and other deep convolutional neural networks, consistently achieve high accuracy on benchmark datasets such as HAM10000 and ISIC, reinforcing the reliability of deep learning in multi-class diagnostic settings [42]. To address the visual similarity between lesion types, hybrid models integrating advanced convolutional blocks with attention mechanisms have also been proposed. These architectures enhance both fine-grained local feature extraction and global contextual representation, resulting in improved precision, recall, and F1-scores across multiple lesion categories [43]. Despite achieving strong classification performance, these studies emphasize the ongoing challenges related to data diversity, interpretability, and computational efficiency. Overall, multi-class deep learning frameworks represent a significant advancement toward comprehensive, scalable, and clinically applicable automated skin cancer diagnostic systems. Nevertheless, the lack of cross-dataset validation in many transfer learning studies raises concerns about the reliability of reported performance metrics.

2.5 Attention Mechanisms in Skin Lesion Analysis

As multi-class classification models developed, researchers increasingly explored attention mechanisms to enhance feature representation and improve diagnostic precision. Attention-based deep learning models were introduced to better focus on clinically significant regions within dermoscopic images, thereby improving the distinction between benign and malignant lesions. Studies comparing multiple deep learning architectures have demonstrated that incorporating attention and model fusion strategies can significantly enhance classification accuracy [44]. In addition, Transformer- inspired architectures have been adapted to medical imaging tasks to overcome the limitations of conventional convolutional operations in capturing global contextual information. Multi-scale cross-attention mechanisms and large-kernel attention techniques have been proposed to simultaneously capture fine-grained local details and long-range spatial dependencies. These hybrid attention-based frameworks have shown superior performance in both segmentation and classification tasks, highlighting the growing importance of attention mechanisms in developing robust and accurate skin lesion analysis systems [45].

2.6 Vision Transformer (ViT) Models

The limited ability of existing convolutional architectures to capture long range relationships has directed growing research attention towards Vision Transformer models for classifying skin lesions. ViT models capture global context through self-attention mechanisms, processing dermoscopic image data at a higher dimensional level and thereby enabling a more complete representation of lesion variability across different types. There are numerous publications analyzing ViT-based classification methods based on using benchmark datasets like ISIC 2019 with high levels of diagnostic confidence versus prior models that struggled with consistently performing well due to image class imbalance and variability [46]. Many researchers are continuing to study the impact of both self-supervised and supervised training on ViT networks to produce features without requiring extensive manual labelling. The use of attention maps for measuring the interpretability of ViT Models can help facilitate determining areas within an image that are clinically relevant; which, in turn, helps build greater trust in the output of the automated process [47]. Furthermore, ViT Models continue to build upon the potential of transformer-based models to produce effective and reliable systems for automated multi-class skin cancer diagnoses. Despite these advances, the high computational cost of attention-based models limits their practical deployment in resource- constrained clinical environments. Figure 3 presents the Vision Transformer architecture, illustrating how input images are divided into patches, embedded with positional information, and processed through multiple transformer blocks using self-attention mechanisms for final classification.

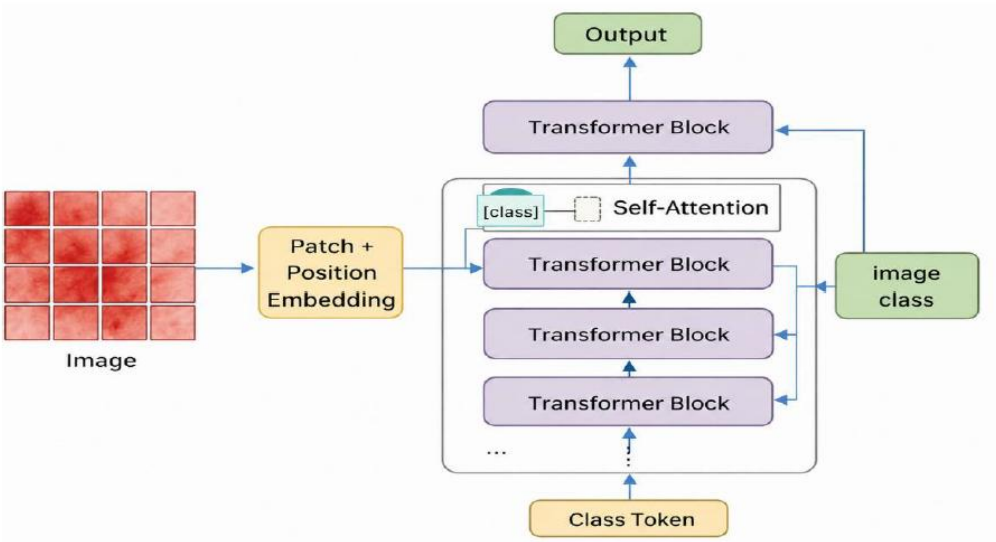

Figure 3 Vision Transformer architecture - patch embedding and self-attention mechanism

2.7 Hybrid Architectures

Recent research has proposed hybrid architectures for further improving the performance of diagnostics [48]. Many of these newly developed methods combine convolutional neural networks (CNN) with frameworks based on the Vision Transformer (ViT) and utilize the strengths of each in order to enhance the early detection of skin cancers [49,50]. Hybrid models that use MetaFormer architecture (focusing specifically on the focal self-attention mechanism) are on par with, or exceed, traditional indicator-based classifiers in terms of their benchmark performance across all classifications specifically ISIC 2019 and HAM10000 (as determined by the classification metrics of accuracy, precision, recall, F1-score) but also possess a lightweight design which helps them to be implemented for clinical purpose in real-time [51-53].

These hybrid architectures have been shown to achieve better performance than many standalone CNNs and ViT models suggesting that the fusion of two complementary approaches can result in superior performance. Concurrently, other research has examined hybrid feature fusion methods that seek to combine handcrafted features (i.e., features that are extracted using conventional methods such as texture analysis using GLCM and wavelet-based descriptors) with deep learning derived features from pre-trained networks (i.e., DenseNet121) using advanced classifier techniques (i.e. Extreme Gradient Boosting (XGBoost) and ensemble classifiers) in order to provide improved classification reliability and robustness[54]. Consider together, these hybrid and fusion-based approaches indicate that fusing multiple generative methods for feature extraction and using architectural fusion techniques can produce systems that perform better and are more generalizable in automated multi-class skin cancer detection [55]. Table 1 provides a comparative summary of the key strengths and limitations of CNN, Vision Transformer, Hybrid, and Explainable AI approaches considered in this section.

Table 1 Comparison between CNN, Vision Transformer, Hybrid, and Explainable AI approaches

Model Type	Strength	Limitation
CNN	High accuracy and efficient feature extraction	Limited interpretability
Vision Transformer (ViT)	Captures global contextual relationships	Requires large-scale datasets
Hybrid (CNN + ViT)	Combines local and global features for improved performance	High computational complexity
Explainable AI (XAI)	Enhances model transparency and trust	Limited clinical reliability

2.8 Explainable AI (XAI) Techniques

Deep learning models have become complicated, resulting in concerns over their interpretability and reliability when used in a clinical setting [56]. With this, techniques from Explainable AI (XAI) have been integrated into the development of skin cancer detection systems [57-59]. While convolutional based and transformer-based architectures have both achieved excellent rates of diagnostic accuracy, their 'black-box' nature has led to a lack of trust by many healthcare providers. For this reason, many researchers have investigated various XAI approaches, such as Grad-CAM, LIME and SHAP, to visualize the regions of dermoscopic or histopathologic images that contribute most heavily to the prediction made by the model [59,60]. Through comparative analysis of MobileNet, DenseNet, XceptionNet and ReXNet-150 architectures, use of high-performing classification models combined with XAI approaches yields improved clinical relevance and user confidence. Lightweight and efficient deep learning models, such as those augmented with Grad-CAM, have shown their ability to highlight regions of lesions while providing real-time inference capability [60]. In addition, hybrid explainable frameworks (for example, using LIME and SHAP) allow for the generation of human-understandable explanations by providing an analysis of low-level feature contributions and high-level feature contributions [61]. The identified improvements demonstrate that integrating XAI with high-performing deep learning models will close the gap of trust between medical practitioners and artificial intelligence systems, thus ensuring the safe and appropriate use of automated multi-class skin cancer diagnostic systems. Since 2020, research interest in skin lesion classification systems has grown considerably [62,63]. The research shows there is an issue with the quality of the datasets used and also how they were created, with a number of duplicate images in the ISIC datasets as well as inconsistency in the splitting into 'training' groups versus 'testing' groups and the lack of identification of which group a particular image belongs to when learning from data. This lack of identification will give an unreliable evaluation of how effective a model would be if it were tested on a dataset [64]. For these reasons, some researchers are recommending that when evaluating datasets for fair benchmarking against a model, it is recommended to use curated datasets and maintain balance [65,66].

2.9 Multimodal Fusion and Domain Adaptation

In addition, researchers have looked at ways to combine both dermoscopic images and patient metadata through advanced fusion techniques using cross-attention mechanisms with the aim of improving the reliability of the embedded dataset and increasing the generalizability of the models built from those datasets [67,68]. Researchers have described using collaborative edge computing techniques for reducing latency and preserving patient data privacy in a clinical setting [69,70]. The issue of domain shift also complicates the evaluation of automated skin lesion analysis models in that the performance of a model when trained on one dataset will be different from that of another dataset used to test that model (i.e., if it is trained on a dataset collected with one device, using a different population, or obtained from a different clinical setting) [71]. Although XAI techniques provide valuable visual explanations, their clinical relevance and reliability in real-world diagnostic settings remain limited and require further investigation.

2.10 Domain Adaptation and Clinical Validation

Few studies have examined how to reduce this performance gap by evaluating the performance of models developed with unsupervised domain adaptation techniques; however, all these studies reported little effect on the performance gap [70]. In addition, recent literature reviews of cutting- edge AI systems stress the increased importance of multimodal fusion, large-scale models, and collaboration between clinicians and AI systems in regard to AI used for skin cancer evaluation, while continuing to note persisting concerns related to data diversity, interpretability, and clinical validation [72-74]. Therefore, several studies have demonstrated that to produce highly reliable, unbiased, and generalizable multi-class skin cancer detection, not only does an AI system need to use high-performance architectures, but also datasets must be thoroughly curated, use appropriate domain adaptation techniques, incorporate multimodal data, and undergo careful assessment in the clinical setting [75]. The lack of standardized multimodal datasets and evaluation protocols continues to hinder the development of reliable and generalizable skin cancer detection systems. While there has been a major advancement for detecting skin cancers using deep learning techniques, there are still many limitations that exist. For example, while many existing models perform well on benchmark datasets (e.g., ISIC, HAM10000), they often perform poorly when tested on practical datasets due to differences in the data that contributes to the way each system works (domain shift) as well as bias from how each dataset was constructed. In addition to the issues described above, most deep learning systems (especially CNNs, and transformer-based systems) have very little interpretability; therefore, making it much harder for clinicians to trust these models when considering using them in practice. Furthermore, while many papers primarily look at accuracy, there are many papers that do not consider how efficiently their model can run in practice or how quickly their model can be deployed (time constraints). All of these problems highlight the urgent need to develop models that are more robust, generalizable, and explainable if they are to be effectively used in a clinical setting [76].

3 Datasets for Skin Cancer Detection

Deep learning algorithms have shown promising capability in accurately diagnosing melanoma. However, their effectiveness is largely determined by the quality and diversity of the datasets used for training and validation. Three large datasets have been widely used for melanoma diagnosis research i.e. the Australian Institute of Dermatology, Hospital da Trindade, and PH2 databases, have been extensively used to evaluate the performance of deep learning approaches [76,77]. These datasets vary in several aspects, including data types, number of images and classes, annotation quality, and their generalizability and susceptibility to bias. Other recent studies have also examined multimodal datasets that combine image data with patient metadata to enhance the robustness and real-world applicability of skin cancer detection systems [78,79]. Table 2 provides the comprehensive list of datasets used to for skin cancer prediction which were in practice.

Table 2 Most used skin cancer datasets

Dataset	Images	Classes	Type	Limitation
ISIC	13,000+	7	Dermoscopic	Class imbalance, duplicates
HAM10000	10,015	7	Dermoscopic	Skin tone bias
PH2	200	3	Dermoscopic	Small dataset
Derm7pt	2,000	7+	Multi-modal	Complex annotations

4 Comparison of Models and Techniques

The ability of machine learning (ML) and deep learning (DL) algorithms to detect skin cancer has been widely evaluated using various datasets and performance metrics. A systematic comparison of both categories highlights their respective strengths, limitations, and suitability for clinical applications. Traditional ML models for skin cancer detection rely on manually engineered features extracted from dermoscopic images. Methods such as Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) have been widely used for binary and multi-class classification of skin lesions. While these approaches achieved reasonable accuracy on smaller datasets such as PH2, they are limited by their dependence on handcrafted features and their inability to scale effectively to large and complex dermoscopic datasets [21]. Deep learning models overcome the limitations of traditional ML by automatically learning hierarchical features directly from raw dermoscopic images. CNN-based architectures such as DenseNet and Xception have demonstrated significantly higher accuracy on large benchmark datasets including ISIC and HAM10000. Transfer learning techniques using pre-trained models such as EfficientNet have further improved performance by leveraging knowledge from large-scale datasets. Most recently, hybrid architectures combining CNN with Vision Transformers (ViT) have achieved state-of-the-art results, reaching up to 95% accuracy on ISIC 2019 by combining local feature ex- traction with global contextual understanding [49].

Table 3 Comparison of Machine Learning and Deep Learning Models

Category	Model	Dataset	Accuracy	Key Feature	Ref
ML	SVM	PH2	79%	Handcrafted features	[21]
ML	KNN	PH2	78%	Simple classifier	[21]
DL	CNN	ISIC	84%	Automatic feature learning	[26]
DL	DenseNet	HAM10000	89%	Deep feature extraction	[36]
DL	Xception	ISIC	91%	High accuracy CNN	[41]
DL	EfficientNet	HAM10000	92%	Lightweight	[37]
DL	CNN + ViT	ISIC 2019	95%	Hybrid architecture	[49]

Table 3 presents a systematic comparison of both ML and DL models reviewed. As shown, DL models consistently outperform traditional ML approaches across all benchmark datasets. While SVM and KNN achieved 78–79% accuracy on PH2, CNN-based models reached 84–92% on ISIC and HAM10000. The hybrid CNN+ViT architecture achieved the highest accuracy of 95% on ISIC 2019, demonstrating the advantage of combining complementary architectural strengths. However, it should be noted that accuracy values vary depending on dataset composition, preprocessing methods, and experimental conditions, and direct comparisons should be interpreted with caution [53]. Based on the findings, the following directions are recommended for future research in multi-class skin cancer detection: There is a need to develop large-scale, diverse, and multi-source dermoscopic datasets that represent different populations, skin tones, and imaging conditions to reduce algorithmic bias and improve model generalizability across varied clinical settings. Advancing multimodal learning frameworks that combine dermoscopic images with clinical metadata such as patient age, sex, and lesion location to enhance diagnostic accuracy and clinical relevance is needed. Improving Explainable AI (XAI) techniques to produce clinically meaningful and reliable explanations that support dermatologists in making accurate and evidence-based diagnostic decisions still a challenge. Designing lightweight and computationally efficient deep learning architectures suitable for real-time deployment in resource-limited healthcare environments is required. Adopting federated learning frameworks and establishing standardized evaluation protocols to enable privacy-preserving collaborative model development across multiple clinical institutions while ensuring regulatory compliance will be required [80].

5 Conclusions

Automated detection of skin cancer has changed dramatically during the last decade, moving from classic techniques that relied upon manually engineered features to complex deep learning systems capable of independently learning visual representations from dermoscopic images. This paper provides various elements of this evolution, covering major developments in CNNs, transfer learning, attention mechanisms, Vision Transformers, hybrid models, and explainable AI methods. Overall, the analysis carried out in this paper illustrates a clear performance trend across all reviewed methods. Traditional ML approaches such as SVM and KNN achieved moderate accuracy of approximately 78–79% on the PH2 dataset due to their reliance on handcrafted features. In contrast, CNN-based models achieved significant accuracy gains of 84–92% on larger datasets such as ISIC and HAM10000 through automatic feature learning. Hybrid CNN+ViT models exhibited the strongest results, achieving approximately 95% accuracy on ISIC 2019 by combining both local and global feature learning. While XAI methods including Grad-CAM, LIME, and SHAP have improved model transparency, further validation is needed to confirm their clinical interpretability and reliability in direct diagnostic settings. While progress has been made in healthcare AI over the past decade, several key challenges remain. Many models produce acceptable results on standard benchmark datasets, but their performance often drops considerably when applied to real clinical data due to differences in imaging conditions and patient demographics. The continuing presence of issues related to class imbalance, annotation inconsistencies, and high computational demands makes practical deployment challenging. In addition, many publicly available datasets lack demographic diversity, meaning that models trained on these datasets may not perform reliably across different patient populations. Addressing these challenges will require coordinated effort across researchers, clinicians, and policymakers. Future work must focus on developing robust, interpretable, and computationally efficient AI systems that deliver consistent performance across different patient populations and varied clinical settings. The continued integration of advanced AI methods with clinical expertise holds strong potential for improving early diagnosis of multi-class skin cancer and supporting better patient outcomes.

Authors' Contributions

All authors contributed equally to the study's conception, design, interpretation, and manuscript preparation. All authors read and approved the final manuscript.

Ethical Approval

Not Applicable

Consent to Participate

Not Applicable

Competing Interests

The authors declare that they have no relevant financial or non-financial interests to disclose.

Data Availability Statement

No data available related to this work

References

[1] A. H. Roky, M. M. Islam, A. M. F. Ahasan, M. S. Mostaq, M. Z. Mahmud, M. N. Amin, and M. A. Mahmud, "Overview of skin cancer types and prevalence rates across continents," *Cancer Pathogens and Therapy*, vol. 3, no. 2, pp. 89–100, 2024, doi: 10.1016/j.cpt.2024.08.002.

[2] A. Naeem, T. Anees, M. Khalil, K. Zahra, R. A. Naqvi, and S. W. Lee, "SNC Net: Skin cancer detection by integrating handcrafted and deep learning-based features using dermoscopy images," *Mathematics*, vol. 12, no. 7, p. 1030, 2024, doi: 10.3390/math12071030.

[3] M. Tahir, A. Naeem, H. Malik, J. Tanveer, R. A. Naqvi, and S. W. Lee, "DSCC Net: Multi-classification deep learning models for diagnosing skin cancer using dermoscopic images," *Cancers*, vol. 15, no. 7, p. 2179, 2023, doi: 10.3390/cancers15072179.

[4] B. Vasudevan, D. Mitra, and V. Deshwal, "Artificial intelligence in dermatology: The shifting paradigms," *Medical Journal Armed Forces India*, 2025, doi: 10.1016/j.mjafi.2025.08.016.

[5] H. Naseri and A. A. Safaei, "Diagnosis and prognosis of melanoma from dermoscopy images using machine learning and deep learning: A systematic literature review," *BMC Cancer*, vol. 25, no. 1, p. 75, 2025, doi: 10.1186/s12885-024-13423-y.
[6] A. M. Zbrzezny and T. Krzywicki, "Artificial intelligence in dermatology: A review of methods, clinical applications, and perspectives," *Applied Sciences*, vol. 15, no. 14, p. 7856, 2025, doi: 10.3390/app15147856.
[7] S. Górecki, A. Tatka, and J. Brusey, "Artificial intelligence and new technologies in melanoma diagnosis: A narrative review," *Cancers*, vol. 17, no. 24, p. 3896, 2025, doi: 10.3390/cancers17243896.
[8] X. Chen, Z. Zhou, H. Ding, H. Zheng, and Y. Ge, "Transforming aesthetic dermatology: The role of artificial intelligence in skin health," *Dermatology and Therapy*, vol. 15, pp. 1999–2013, 2025, doi: 10.1007/s13555-025-01459-2.
[9] S. S. Han, S. I. Cho, S. E. Chang, S. H. Kim, and J. I. Na, "Validation of an AI for skin diseases in Korea and global usage statistics," *medRxiv*, 2025, doi: 10.1101/2025.01.27.25321219.
[10] N. Nadour, T. Duguet, S. Zahedi, H. Figoni, and R. Liard, "Diagnostic accuracy of artificial intelligence compared to family physicians and dermatologists for skin conditions: A systematic review and meta-analysis," *BMC Primary Care*, vol. 26, p. 384, 2025, doi: 10.1186/s12875-025-03073-9.
[11] N. Verma, Ranvijay, and D. K. Yadav, "A comprehensive review on step-based skin cancer detection using machine learning and deep learning methods," *Archives of Computational Methods in Engineering*, vol. 32, pp. 4359–4412, 2025, doi: 10.1007/s11831-025-10275-y.
[12] G. Narmadha and M. Sivasakthi, "An investigation on the detection and classification of skin cancer using deep learning," in *Artificial Intelligence and Technology: Systems Management, Decisions and Control for Sustainability in the Digital Age*, pp. 567–576, 2026.
[13] Y. K. Ahmed and A. N. A. Naji, "Smart feature extraction using deep learning for early diagnosis of chronic diseases," *Network Modeling Analysis in Health Informatics and Bioinformatics*, vol. 14, p. 140, 2025, doi: 10.1007/s13721-025-00641-y.
[14] A. Alrabai, A. Echtioui, and F. Kallel, "Explainable transfer learning for skin cancer diagnosis," *Network Modeling Analysis in Health Informatics and Bioinformatics*, vol. 15, no. 1, p. 30, 2026, doi: 10.1007/s13721-025-00709-9.
[15] O. Attallah, "Skin-CAD: Explainable deep learning classification of skin cancer from dermoscopic images by feature selection of dual high-level CNN features and transfer learning," *Computers in Biology and Medicine*, vol. 178, p. 108798, 2024, doi: 10.1016/j.compbiomed.2024.108798.
[16] H. AbuAlkebash, R. A. A. Saleh, and H. M. Ertunc, "Explainable depth-wise and channel-wise fusion models for multi-class skin lesion classification," *PLoS One*, vol. 21, no. 1, p. e0340901, 2026, doi: 10.1371/journal.pone.0340901.
[17] A. A. Noor, A. Manzoor, M. D. M. Qureshi, M. A. Qureshi, and W. Rashwan, "Unveiling explainable AI in healthcare: Current trends, challenges, and future directions," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2025, doi: 10.1002/widm.70018.
[18] O. Uwishema, M. Ghezzawi, N. Charbel, S. Alawieh, S. Roy, M. Wojtara, C. M. Hakayuwa, I. K. Ja'afar, and M. Prasad, "Diagnostic performance of artificial intelligence for dermatological conditions: A systematic review," *International Journal of Emergency Medicine*, vol. 18, p. 172, 2025, doi: 10.1186/s12245-025-00975-4.
[19] M. S. Yaswanth and C. M. K. Reddy, "Unresolved challenges and future directions in AI-based skin cancer detection," in *Proc. IEEE ICECA*, 2025, doi: 10.1109/ICECA66444.2025.11382676.
[20] T. Majtner, S. Yildirim-Yayilgan, and J. Y. Hardeberg, "Combining deep learning and hand-crafted features for skin lesion classification," in *Proc. 6th Int. Conf. Image Processing Theory, Tools and Applications (IPTA)*, 2016, pp. 1–6, doi: 10.1109/IPTA.2016.7821017.
[21] E. M. Senan and M. E. Jadhav, "Diagnosis of dermoscopy images for the detection of skin lesions using SVM and KNN," in *Proc. 3rd Int. Conf. Sustainable Computing (SUSCOM 2021)*, Singapore: Springer, 2022, pp. 125–134.
[22] M. A. Kassem, K. M. Hosny, R. Damaševičius, and M. M. Eltoukhy, "Machine learning and deep learning methods for skin lesion classification and diagnosis: A systematic review," *Diagnostics*, vol. 11, no. 8, p. 1390, 2021, doi: 10.3390/diagnostics11081390.
[23] T. R. S. Fernandes, A. S. Teles, J. R. N. Fernandes, L. D. B. Lima, D. L. Sousa, and R. C. Soares, "External validation of AI models for skin diseases: A systematic review," *IEEE Access*, 2025, doi: 10.1109/ACCESS.2025.3584904.
[24] M. Naqvi, S. Q. Gilani, T. Syed, O. Marques, and H.-C. Kim, "Skin cancer detection using deep learning—A review," *Diagnostics*, vol. 13, no. 11, p. 1911, 2023, doi: 10.3390/diagnostics13111911.
[25] M. A. Albahar, "Skin lesion classification using convolutional neural network with novel regularizer," *IEEE Access*, vol. 7, pp. 38306–38313, 2019, doi: 10.1109/ACCESS.2019.2906241.
[26] R. Tanna and T. Sharma, "Binary classification of melanoma skin cancer using SVM and CNN," in *Proc. Int. Conf. Artificial Intelligence and Machine Vision (AIMV)*, 2021, pp. 1–4, doi: 10.1109/AIMV53313.2021.9670894.
[27] N. M. Rashad, N. M. M. Abdelnapi, A. F. Seddik, and M. A. Sayedelahl, "Automating skin cancer screening: A

deep learning approach," *Journal of Engineering and Applied Science*, vol. 72, p. 6, 2025, doi: 10.1186/s44147-024-00573-w.
[28] G. Durga Devi and D. M. D. Preethi, "Design and analysis of melanoma skin cancer using deep learning and machine learning," in *Proc. IEEE ICDICI*, 2025, doi: 10.1109/ICDICI66477.2025.11134986.
[29] K. Wanjari and P. Verma, "A review on the applications of machine learning and deep learning techniques for skin cancer detection," in *Proc. IEEE ICSADL*, 2025, doi: 10.1109/ICSADL65848.2025.10933289.
[30] S. A. H. Shah, S. T. H. Shah, R. Khaled, A. Buccoliero, S. B. H. Shah, A. Di Terlizzi, G. Di Benedetto, and M. A. Deriu, "Explainable AI-based skin cancer detection using CNN, particle swarm optimization and machine learning," *Journal of Imaging*, vol. 10, no. 12, p. 332, 2024, doi: 10.3390/jimaging10120332.
[31] Z. M. Gohil and M. B. Desai, "Revolutionizing dermatology: A comprehensive survey of AI-enhanced early skin cancer diagnosis," *Archives of Computational Methods in Engineering*, vol. 31, pp. 4521–4531, 2024, doi: 10.1007/s11831-024-10121-7.
[32] H. A. Haenssle, C. Fink, R. Schneiderbauer, F. Toberer, T. Buhl, A. Blum, A. Kalloo, A. B. H. Hassen, L. Thomas, A. Enk, and L. Uhlmann, "Man against machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists," *Annals of Oncology*, vol. 29, no. 8, pp. 1836–1842, 2018, doi: 10.1093/annonc/mdy166.
[33] H. Feng and L. Zhang, "EfficientNet-based skin cancer classification and recognition," in *Proc. SPIE*, vol. 13800, p. 138001Q, 2025, doi: 10.1117/12.3076682.
[34] A. Mahbod, G. Schaefer, C. Wang, G. Dorffner, R. Ecker, and I. Ellinger, "Transfer learning using a multi-scale and multi-network ensemble for skin lesion classification," *Computer Methods and Programs in Biomedicine*, vol. 193, p. 105475, 2020, doi: 10.1016/j.cmpb.2020.105475.
[35] M. Zakariah, M. Al-Razgan, and T. Alfakih, "Skin cancer detection with MobileNet-based transfer learning and MixNets for enhanced diagnosis," *Neural Computing and Applications*, vol. 36, pp. 21383–21413, 2024, doi: 10.1007/s00521-024-10227-w.
[36] A. Bello, S.-C. Ng, and M.-F. Leung, "Skin cancer classification using fine-tuned transfer learning of DenseNet-121," *Applied Sciences*, vol. 14, no. 17, p. 7707, 2024, doi: 10.3390/app14177707.
[37] V. Anand, S. Gupta, S. R. Nayak, D. Koundal, D. Prakash, and K. D. Verma, "An automated deep learning model for classification of skin disease using dermoscopy images: A comprehensive study," *Multimedia Tools and Applications*, vol. 81, no. 26, pp. 37379–37401, 2022, doi: 10.1007/s11042-021-11628-y.
[38] M. Ennab and H. Mcheick, "Advancing AI interpretability in medical imaging: A comparative analysis of pixel-level interpretability and Grad-CAM models," *Machine Learning and Knowledge Extraction*, vol. 7, no. 1, p. 12, 2025, doi: 10.3390/make7010012.
[39] S. Sharma, A. Sharma, A. Verma, N. K. Fuloria, and A. Sharma, "Enhancing skin cancer detection with transfer learning and deep learning," in *AI and IoT-Enhanced Skin Cancer Detection and Care*, pp. 233–272, 2025, doi: 10.2174/9798898119521250101.
[40] C. Revanth, T. Balaji, N. K. Kumar, D. S. K. Goud, G. S. Haritha, and B. Praveen, "Optimizing skin cancer detection using MobileNet," in *Proc. IEEE AIMLA*, 2025, doi: 10.1109/AIMLA63829.2025.11040509.
[41] M. Badr, A. Elkasaby, M. Alrahmawy, and S. El-Metwally, "A multi-model deep learning architecture for diagnosing multi-class skin diseases," *Journal of Imaging Informatics in Medicine*, vol. 38, no. 3, pp. 1776–1795, 2025, doi: 10.1007/s10278-024-01300-w.
[42] P. Dubey, P. Bande, R. Shukla, S. S. Kshatri, M. Nayak, and R. Sharma, "AI-powered dermatology: Revolutionizing skin cancer detection through machine learning algorithms," in *Proc. Int. Conf. Cybernation and Computation (CYBERCOM)*, IEEE, 2024, doi: 10.1109/CYBERCOM63683.2024.10803244.
[43] B. Ozdemir and I. Pacal, "A robust deep learning framework for multiclass skin cancer classification," *Scientific Reports*, vol. 15, no. 1, p. 4938, 2025, doi: 10.1038/s41598-025-89230-7.
[44] H. Chelabi, B. Chikhaoui, A. Kermi, and M. T. Khadir, "Classification of skin lesion images based on attention mechanism," in *Proc. IEEE/ACS 22nd Int. Conf. Computer Systems and Applications (AICCSA)*, 2025, pp. 1–7, doi: 10.1109/AICCSA66935.2025.11315191.
[45] S. Liu, Z. Zhuang, Y. Zheng, and S. Kolmaniče, "A VAN-based multi-scale cross-attention mechanism for skin lesion segmentation network," *IEEE Access*, vol. 11, pp. 81953–81964, 2023, doi: 10.1109/ACCESS.2023.3298826.
[46] A. R. Nath and S. Bhattacharjee, "A novel vision transformer based approach for skin lesion classification: A deep learning framework," in *Proc. Int. Conf. Quantum Photonics, Artificial Intelligence, and Networking (QPAIN)*, 2025, pp. 1–5, doi: 10.1109/QPAIN66474.2025.11172224.
[47] R. Abdel Mawgoud and C. Posch, "Visual explainability of 250 skin diseases viewed through the eyes of an AI-based, self-supervised vision transformer—A clinical perspective," *JEADV Clinical Practice*, vol. 4, no. 1, pp. 145–155, 2025, doi: 10.1002/jvc2.580.
[48] S. P. Jeong, J. Liu, and S. J. Fong, "Bridging deep learning in skincare and dermatology: Opportunities for hybrid models and early skin cancer prediction," *International Journal of Contents*, vol. 21, no. 2, p. 24, 2025, doi: 10.5392/ijoc.2025.21.2.024.
[49] I. Pacal, B. Ozdemir, J. Zeynalov, H. Gasimov, and N. Pacal, "A novel CNN-ViT-based deep learning model for

early skin cancer diagnosis," *Biomedical Signal Processing and Control*, vol. 104, p. 107627, 2025, doi: 10.1016/j.bspc.2025.107627.
[50] A. El Mahdi, R. Abla, E. Assia, S. Abdelouahed, and A. Abdellah, "Fusion of Vision Transformer and VGG features for enhanced skin lesion classification," in *Proc. ICCSC 2025*, 2025, doi: 10.1109/ICCSC66714.2025.11135399.
[51] M. Z. Hasan and F. Y. Rifat, "Hybrid ensemble of segmentation-assisted classification and GBDT for skin cancer detection," *arXiv preprint arXiv:2506.03420*, 2025, doi: 10.48550/arXiv.2506.03420.
[52] K. Palaniswamy, M. MeenakshiSundaram, S. Arige, and T. B. Krishna, "Multimodal deep learning for skin cancer diagnosis using images and metadata," in *Proc. IEEE CICT*, 2025, doi: 10.1109/CICT67193.2025.11398859.
[53] P. G. Dhange and P. William, "Exploring deep neural architectures for skin lesion detection and classification," in *Proc. Int. Conf. Computational Intelligence and Knowledge Economy (ICCIKE)*, IEEE, 2025, pp. 1–6, doi: 10.1109/ICCIKE67021.2025.11318236.
[54] B. Soundarya and C. Poongodi, "A novel hybrid feature fusion approach using hand-crafted features with transfer learning model for enhanced skin cancer classification," *Computers in Biology and Medicine*, vol. 190, p. 110104, 2025, doi: 10.1016/j.compbiomed.2025.110104.
[55] K. Sreekala, N. Rajkumar, R. Sugumar, K. V. D. Sagar, R. Shobarani, K. P. Krishnamoorthy, A. K. Saini, H. Palivela, and A. Yeshitla, "Skin diseases classification using a hybrid AI-based localization approach," *Computational Intelligence and Neuroscience*, vol. 2022, p. 6138490, 2022, doi: 10.1155/2022/6138490.
[56] H. H. Aras and N. Doğan, "A comparative analysis for skin cancer detection by using explainable deep learning," *Neural Computing and Applications*, vol. 38, no. 4, p. 50, 2026.
[57] M. B. Jan, M. Rashid, R. Vavekanand, and V. Singh, "Integrating explainable AI for skin lesion classifications: A systematic literature review," *SMHS Journal*, vol. 2, no. 1, 2023, doi: 10.48185/smhs.v2i1.1422.
[58] D. Khurana, R. Kaur, R. K. Roul, and S. Batra, "Explainable artificial intelligence in cancer care: A domain-wise review of adoption, challenges and opportunities," *Expert Systems*, 2026, doi: 10.1111/exsy.70238.
[59] Y. He, J. Zheng, and E. Zou, "SHARPEN-CAM: Efficient hierarchical SHAP-based visual explanation for deep CNNs," *Multimedia Systems*, vol. 32, p. 3, 2026.
[60] S. Suara, A. Jha, P. Sinha, and A. A. Sekh, "Is Grad-CAM explainable in medical images?" in *Computer Vision and Image Processing (CVIP 2023)*, pp. 124–135, 2024.
[61] A. W. Lubis, A. G. Putrada, and A. Q. Istiqfarri, "ReXNet-150 and Grad-CAM for efficient and explainable melanoma skin cancer detection," in *Proc. 5th Int. Conf. Intelligent Cybernetics Technology and Applications (ICICyTA)*, 2025, pp. 137–142, doi: 10.1109/ICICyTA68677.2025.11362836.
[62] C. C. Ukwuoma, D. Cai, E. O. Eziefuna, A. Oluwasanmi, S. F. Abdi, G. W. Muoka, D. Thomas, and K. Sarpong, "Enhancing histopathological medical image classification for early cancer diagnosis using deep learning and explainable AI—LIME and SHAP," *Biomedical Signal Processing and Control*, vol. 100, p. 107014, 2025, doi: 10.1016/j.bspc.2024.107014.
[63] Z. S. A. Wahid, Z. C. Oleiwi, and R. F. Kadhem, "An analytical and comparative study of modern intelligent models for skin cancer diagnosis," *Journal of Quantum Computational Science Modeling*, vol. 18, p. 12671, 2026, doi: 10.29304/jqcsm.2026.18.12671.
[64] J. W. Tjiu and C. F. Lu, "Equity and generalizability of artificial intelligence for skin-lesion diagnosis using clinical, dermoscopic, and smartphone images: A systematic review and meta-analysis," *Medicina*, vol. 61, no. 12, p. 2186, 2025, doi: 10.3390/medicina61122186.
[65] B. Cassidy, C. Kendrick, A. Brodzicki, J. Jaworek-Korjakowska, and M. H. Yap, "Analysis of the ISIC image datasets: Usage, benchmarks and recommendations," *Medical Image Analysis*, vol. 75, p. 102305, 2022, doi: 10.1016/j.media.2021.102305.
[66] D. Ajabani, Z. A. Shaikh, A. Yousef, K. Ali, and M. A. Albahar, "Enhancing skin lesion classification with human baseline comparison," *PeerJ Computer Science*, vol. 11, p. e2795, 2025, doi: 10.7717/peerj-cs.2795.
[67] A. M. Shabu, M. S. Ansari, and A. Aslam, "A generative AI approach for reducing skin tone bias in skin cancer classification," *arXiv preprint arXiv:2602.14356*, 2026, doi: 10.48550/arXiv.2602.14356.
[68] A. Solak, "HCHS-Net: A multimodal handcrafted feature and metadata framework for interpretable skin lesion classification," *Biomimetics*, vol. 11, no. 2, p. 154, 2026, doi: 10.3390/biomimetics11020154.
[69] N.-Y. Tran-Van and K.-H. Le, "A multimodal skin lesion classification through cross-attention fusion and collaborative edge computing," *Computerized Medical Imaging and Graphics*, vol. 124, p. 102588, 2025, doi: 10.1016/j.compmedimag.2025.102588.
[70] M. M. Yaqoob, M. Alsulami, M. A. Khan, and coauthors, "Symmetry in privacy-based health-care: A review of skin cancer detection using federated learning," *Symmetry*, vol. 15, no. 7, p. 1369, 2023, doi: 10.3390/sym15071369.
[71] S. Chamarthi, K. Fogelberg, T. J. Brinker, and J. Niebling, "Mitigating the influence of domain shift in skin lesion classification: A benchmark study of unsupervised domain adaptation methods," *Informatics in Medicine Unlocked*, vol. 44, p. 101430, 2024, doi: 10.1016/j.imu.2023.101430.
[72] N. Alipour, T. Burke, and J. Courtney, "Skin type diversity in skin lesion datasets: A review," *Current Dermatology Reports*, vol. 13, pp. 198–210, 2024, doi: 10.1007/s13671-024-00440-0.

[73] Y. Diao, X. Chen, Z. Huang, Q. Tan, M. Yang, H. Yu, Y. Xu, and X. Hu, "Cutting-edge AI technologies in skin cancer applications," *Cancer Letters*, p. 218256, 2026, doi: 10.1016/j.canlet.2026.218256.

[74] N. Veeramani and P. Jayaraman, "A promising AI-based super-resolution image reconstruction technique for early diagnosis of skin cancer," *Scientific Reports*, vol. 15, p. 5084, 2025, doi: 10.1038/s41598-025-89693-8.

[75] N. M. Eshwarappa, H. Baghban, C. H. Hsu, P. Y. Hsu, R. H. Hwang, and M. Y. Chen, "Communication-efficient federated learning for medical image classification," *Journal of Cloud Computing*, vol. 14, p. 44, 2025, doi: 10.1186/s13677-025-00734-z.

[76] C. Supriyanto, A. Salam, J. Zeniarja, D. W. Utomo, I. N. Dewi, C. Paramita, A. Wijaya, and N. Z. M. Safar, "A bibliometric review of deep learning approaches in skin cancer research," *Computation*, vol. 13, no. 3, p. 78, 2025, doi: 10.3390/computation13030078.

[77] M. Alsaidi, M. T. Jan, A. Altaher, H. Zhuang, and X. Zhu, "Tackling class imbalance in dermoscopic image classification using data augmentation and GAN," *Multimedia Tools and Applications*, vol. 83, pp. 49121–49147, 2024, doi: 10.1007/s11042-023-17067-1.

[78] V. Sydorskyi, I. Krashenyi, and O. Yakubenko, "Multimodal system for skin cancer detection," *arXiv preprint arXiv:2601.14822*, 2026, doi: 10.48550/arXiv.2601.14822.

[79] J. Kawahara, S. Daneshvar, G. Argenziano, and G. Hamarneh, "Seven-point checklist and skin lesion classification using multitask multimodal neural networks," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 2, pp. 538–546, 2019, doi: 10.1109/JBHI.2018.2824327.

[80] Pulaski, H., Harrison, S.A., Mehta, S.S., Sanyal, A.J., Vitali, M.C., Manigat, L.C., Hou, H., Madasu Christudoss, S.P., Hoffman, S.M., Stanford-Moore, A. and Egger, R., 2025. Clinical validation of an AI-based pathology tool for scoring of metabolic dysfunction-associated steatohepatitis. Nature Medicine, 31(1), pp.315-322., doi: 10.1038/s41591-024-03301-2.