SSRG International Journal of Computer Science and Engineering
ISSN: 2348–8387 / https://doi.org/10.14445/23488387/IJCSE-V13I5P101

Volume 13 Issue 5, 1-8, May 2026


*Review Article*

# Deep Learning for Melanoma Detection: A Concise Review of Datasets, Architectures, and Clinical Challenges

Mayuri Patil[1], Suraj Redekar[2]

[1,2]*Computer Science and Engineering, Ashokarao Mane Group of Institutions, Vathar, Maharashtra, India.*

[1]*Corresponding Author : mayuripatiluttre@gmail.com)*

Received: 20 March 2026 Revised: 27 April 2026 Accepted: 12 May 2026 Published: 26 May 2026

***Abstract*** *- Melanoma continues to be one of the most prevalent causes of mortality cases involving skin cancers. Therefore, early diagnosis is vital to help increase patient survival rates. This review article aims to provide in-depth information on deep learning approaches applied to melanoma detection and analyze papers on this subject written between 2016 and 2025. Also, the paper will highlight publicly available databases, novel architectures, and existing difficulties associated with their application. In particular, it will explore the features of three popular datasets (ISIC, HAM10000, and PH2) and analyze their strengths and weaknesses. The focus will be on such factors as the class imbalance problem and the underrepresentation of various demographics. In addition, different architectures used for melanoma detection will be compared, namely VGG, ResNet, Inception, and novel Vision Transformer networks based on their efficiency and capacity to classify melanoma. Finally, major challenges faced during implementation, including overfitting of models, difficulty in their generalization to other populations, and the problem of black box methods, will be considered. At the same time, potential future directions of research, such as the application of explainable artificial intelligence, federated learning techniques, and the use of various demographic groups as training data, will be presented.*

***Keywords*** *- Melanoma, Deep Learning, Explainable AI, Clinical Data, Skin Cancer.*

## 1. Introduction

Melanoma is one of the deadliest and most aggressive types of skin cancer that poses a severe clinical challenge worldwide because of its tendency to metastasize quickly [1]. The ability of a tumor to invade and metastasize to distant organs, which adversely affects patient outcome and increases the mortality rate, is the reason why it has such clinical significance [2], [3]. Early-stage melanomas tend to be thin and highly susceptible to curative therapies, resulting in high survival rates and a low mortality rate [4]. Methods for detecting melanomas include skin self-examination, skin examination by a doctor using dermatoscopy, and total body photography. These diagnostic techniques have been improved through recent advancements in deep learning and Artificial Intelligence (AI) technologies [4], [5], and [6]. This combination makes it easier for patients at risk to screen for early-stage melanomas, thus decreasing their mortality rate [6], [7].

The following represent some of the limitations associated with the traditional techniques used in diagnosing melanomas. While there is evidence that dermoscopy increases sensitivity and specificity, the traditional method of diagnosing melanomas by visual examination with the naked eye is highly subjective and highly unreliable since they require factors such as the size of the lesion, blurring, skin type, and image quality to distinguish benign lesions from melanoma [8], [9]. Traditional techniques are also known to usually only identify melanomas at a much later stage, reducing the efficacy of treatment and negatively affecting patient prognosis [5]. Other traditional techniques, such as Tzanck smear, potassium hydroxide preparations, and Woods lamp examinations, have high convenience and availability; however, their ability to accurately diagnose melanoma is highly limited [10]. Non-invasive imaging techniques such as electrical impedance spectroscopy, confocal microscopy, and total body photography show great potential; however, they still lack sufficient validation and standardization required for them to be implemented in practice [9], [11]. They are also highly dependent on the skills of the individual using the instrument. These problems highlight the need for an alternative technique that can enhance and complement the clinical examination process in diagnosing melanomas.

The current critical review addresses the utilization of deep learning for the purpose of detecting melanoma through a comprehensive analysis of the literature on peer-reviewed articles within the timeline of 2016 to 2025 from such sources as PubMed, IEEE Xplore, Scopus, and Web of Science. The annotated bibliographies include the following sections: Annotated Bibliography on Melanoma Detection Datasets Using Deep Learning: M.P. includes: (1) publicly available databases (ISIC, HAM10000, PH2); detailed information about the database, its advantages and disadvantages; (2) the review of deep learning algorithms (Vision Transformer vs. VGG, ResNet, and Inception); comparison of their performance; and (3) methods for evaluation of clinical application (diversity of the training dataset; generalizability to other patient population; interpretability of model outputs and decision making). The main sections in the article are: The importance of melanoma, its detection, and diagnosis are discussed in section one. Section two covers several properties of the publicly available databases. Section three reviews the performance of deep learning algorithms. Section four covers some problems of overfitting, generalizability, and interpretability of the results while working with those databases. Section five discusses future directions of federated learning, XAI, and clinical application research. Conclusion and recommendations for further investigation are provided in Section 6.

There has been increasing research on deep learning models for skin lesion classification. However, current literature on this topic focuses on performance measures and does not sufficiently examine the factors that hinder the use of such models in clinical practice. Critical limitations of the current research include: (1) a lack of comparative assessment between the model's generalizability and the constraints of the dataset used; (2) a lack of systematic reporting on the process of external validation; and (3) a limited investigation into the issues associated with ethics, regulation, and model explainability. Through the exploration of the bias present in datasets, model architecture, and clinical deployment, this review aims to address these gaps.

Three aspects distinguish this review from previous literature on the same topic: (a) a systematic comparison between CNNs and Vision Transformers, both in terms of their accuracy and data efficiency; (b) an analysis of the limitations of the ISIC, HAM10000, and PH2 datasets, apart from class imbalance; and (c) the proposal of an evaluation framework (see Table 2, Section 4).

## 2. Public Datasets

Melanoma detection processes employing machine learning algorithms have extensively utilized the three open-access data sets (PH2, HAM10000, and ISIC), especially in cases where deep learning algorithms have been applied. The ISIC data set, containing a huge number of dermoscopy images from across the globe, has been instrumental in generating reliable algorithms for melanoma segmentation and classification. The HAM10000 data set comprises more than 10,000 annotated images of seven types of skin lesions, one of them being melanoma, hence contributing towards the resolution of complex multi-classification scenarios.

**Table 1. Dataset Summary**

| Dataset | Category | Attributes / Features | Description |
|---|---|---|---|
| ISIC | Dermoscopic images of skin lesions | Several kinds of skin lesions, such as benign lesions and melanoma. Color variation, texture patterns, lesion borders, asymmetry, and size retrieved from dermoscopic pictures are some of the features. | A sizable and varied collection of dermoscopic pictures for melanoma diagnosis is offered by the International Skin Imaging Collaboration (ISIC) dataset. It is frequently used to train and evaluate deep learning and machine learning models for tasks involving the classification and segmentation of skin lesions [14], [15] |
| HAM10000 | Dermatoscopic images of pigmented skin lesions | Seven categories of pigmented skin lesions, including melanoma. Attributes include color distribution, texture, shape descriptors, lesion boundaries, and structural patterns. | HAM10000 is a large, publicly available dataset containing over 10,000 dermatoscopic images. It is extensively used in machine learning research for skin lesion classification and melanoma detection, supporting deep learning-based automated diagnostic frameworks.[16], [17] |
| PH2 | Dermoscopic images of melanocytic lesions | Three lesion categories: common nevi, atypical nevi, and melanoma. Features include lesion color, texture, diameter, asymmetry, and clinically validated annotations. | PH2 is a curated dermatology dataset designed for melanoma research. It serves as a benchmark dataset for evaluating machine learning algorithms in skin lesion segmentation and classification due to its high-quality clinical annotations. [14], [18] |

On the other hand, the PH2 data set is highly specific since it consists of annotated images of melanocytic lesions, making it possible for the user to carry out a comprehensive analysis of the lesions' segmentation and classification. The following are some features usually extracted from the data sets in order to ensure that the algorithm distinguishes between benign and malignant lesions: color, texture, shape, border properties, and size. Consequently, these data sets have led to the development of innovative AI-based models, thereby improving the accuracy of AI-based automatic melanoma diagnosis and early detection systems, both of which are vital in minimizing melanoma fatalities [14], [16], and [18].

PubMed, IEEE Xplore, Scopus, and Web of Science were employed for a comprehensive literature search from January 2016 to March 2025. Some of the keywords that have been utilized are "melanoma detection," "deep learning," "CNN," "Vision Transformer," "dermoscopy database," "deployment in clinical practice," and "explainable AI." Peer-reviewed studies, use of ISIC, HAM10000, and/or PH2 databases, and publication of classification accuracy (AUC, sensitivity, and specificity) are the inclusion criteria. The exclusion criteria include the following: (a) non-English manuscripts; (b) studies without quantifiable outcomes; and (c) reviews not focused on melanoma. Of the total 127 articles identified initially, only 38 have been included in the review after full-text examination of 52 articles.

The ISIC, HAM10000, and PH2 databases each possess unique qualities that can be valuable in the investigation of melanoma detection, but they also have several drawbacks. Even though the ISIC database features a large amount of images of skin lesions, labeled according to their type, allowing for effective model training and generalization of various clinical presentations, it suffers from the problems of class imbalance and varying quality of images, resulting in unpredictable performance of the model [19], [20]. The HAM10000 dataset features various images of dermatological cases with dermatoscopic details that are specifically labeled for seven skin lesion types, among which there is melanoma. The images can be used for multi-task categorization studies due to their rich description of color, texture, and morphology features. The unequal representation of each type of image inside each category makes the search for images depicting rare types of skin lesions rather difficult [21], [22]. The PH2 dataset is too small to train deep learning models effectively without risking overfitting. Nevertheless, its rich annotation of melanocytic lesions allows for segmentation and classification research. However, the small size of the PH2 database and the limited variation of its samples make it insufficient to reflect real-world conditions encountered in routine practice [14], [15]. In addition, most of these databases fail to represent individuals with a wide variety of skin colors and demographics, which might hinder the applicability of AI models to various groups. While the above datasets have considerably improved algorithms for the detection of melanoma, scientists continue working to resolve the present problems related to class imbalances, small sample size of rare skin lesions, varying image quality, and demographic constraints through the use of data augmentation and multimodality techniques [19], [21], [22].

Images from fairer skin types (Fitzpatrick Skin Type I-III) dominate ISIC, HAM10000, and PH2 datasets. The proportion of images for types V-VI is under 5%. Thus, the ability of the model to diagnose melanoma on different skin types, including acral lentiginous melanoma, is lowered. Model training may lead to systematic bias in diagnosing diseases without clear fairness metrics like equalized odds.

Data augmentation and sampling methods have been widely used to address the problem of underrepresentation of minority groups and class imbalance in melanoma datasets. Rotations, scales, flips, and generation of synthetic images through methods like SMOTE and GANs have been cited by authors as some methods that can be used to increase the number of samples for the minority class and enhance the diversity of the dataset, in addition to increasing class representation [23], [24]. The imbalance between classes is overcome by giving preference to the minority classes in model training through class weighting [25]. Sampling strategies like undersampling the dominant class and oversampling the minority class help to balance the dataset [26], [27]. Unlabeled data and pre-trained knowledge have been utilized in semi-supervised learning and transfer learning models, respectively, to counteract the problems associated with insufficient data and imbalances [28].

## 3. Deep Learning Architectures

Popular CNN architectures for melanoma detection research include VGG, ResNet, and Inception; these architectures offer different strengths compared to others. For instance, the VGG architecture (including VGG by 16s and VGG by 19s) has proved very effective in sensitivity and accuracy while classifying skin lesions, despite having a shallow depth for research purposes. [29], [30]. Due to the residual learning ability that allows one to train deep networks without facing problems associated with decay, the ResNet architecture (e.g., ResNet50) is commonly used. It has been found to be helpful in melanoma classification through databases such as ISIC [30], [31]. Inception architecture (e.g., Inception-v3 and a combination of Inception-ResNet) is used for its multi-scale feature extraction capability, especially when combined with transfer learning techniques [29], [32]. Hybridization techniques are often applied to the CNN models to enhance their accuracy. The main factors behind the frequent application of the VGG model in melanoma detection research include its outstanding performance and simplicity, support from the ResNet model in deep network training, and multi-scale context analysis offered by the Inception model [25].

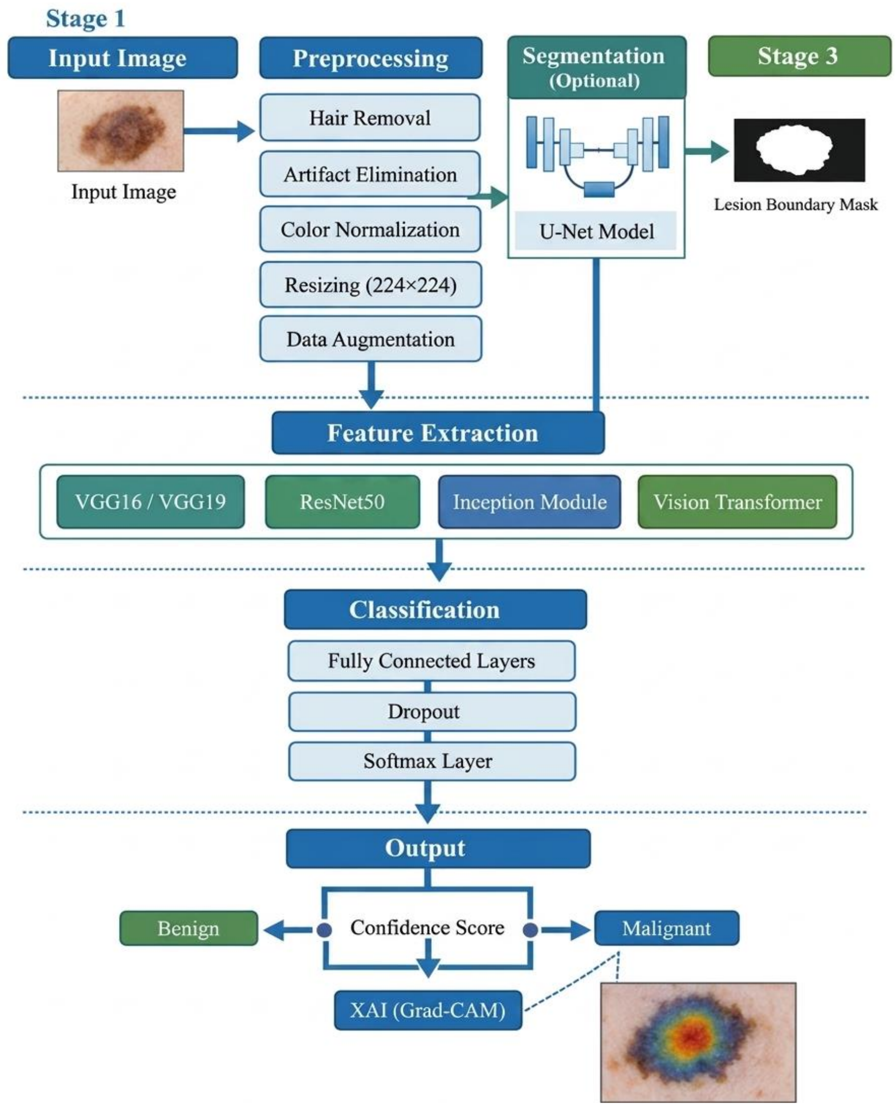


**Fig. 1 Typical deep learning pipeline for melanoma detection**

Popular CNN architectures used for studying melanoma detection include VGG, ResNet, and Inception, and these CNN architectures have distinctive advantages over other CNN architectures. For example, VGG architecture (VGG 16s and VGG 19s) has shown a significant level of effectiveness in terms of sensitivity and accuracy when it comes to skin lesion classification despite the architecture being shallow in depth for research purposes [29], [30]. ResNet architecture (ResNet50), which enables residual learning and helps one to train deep networks without any decay issues, is frequently used due to the architecture's ability to help researchers classify melanomas based on databases like ISIC [30], [31]. Inception architecture (Inception-v3 and Inception-ResNet) is chosen for multi-scale feature extraction through the use of transfer learning [29], [32]. CNN models' hybridization is employed in order to improve the accuracy levels. Key reasons explaining why the VGG architecture has frequently been used in studying melanoma detection include the VGG architecture's efficiency and simplicity, the ResNet architecture's support in training deep networks, and multi-scale context analysis through the Inception architecture [25].

Figure 1 shows an end-to-end system architecture for the melanoma classification of dermoscopic images through deep learning techniques, consisting of ordered phases that have been specifically defined. The dermoscopic image first goes through the preparation phase that consists of hair removal, artifact removal, color normalization, downsizing to 224 × 224 pixels, and data augmentation for improved robustness. In the optional phase of lesion segmentation, the lesion area is delineated, and the border mask is created using a U-Net architecture. Discriminatory features from the prepared image/lesion mask can be acquired through versatile deep learning structures such as VGG16/19, ResNet50 with skip connection, Inception models, or a Vision Transformer with patch embedding and attention models. These features are fed into a classification model, which includes dense layers, dropout, and a softmax layer for aiding in decision-making. The final stage ensures interpretability and relevancy through the use of an output model that provides a binary classification (benign/malignant) along with a probability score and an explanation of the findings through an explainable artificial intelligence system using Grad-CAM visualizations.

The Graph Neural Network (GNN) approach represents skin lesions as graphs of superpixels, incorporating topological connections beneficial for detecting irregular borders, unlike CNN and Vision Transformer approaches. The dependence on vast datasets is reduced by utilizing Self-Supervised Learning (SSL), which has shown potential in training on unlabelled dermoscopic images. Preliminary studies indicate that SSL may reach up to 85% of supervised learning accuracy using only 10% of labeled data. However, although promising for generalization among underrepresented skin colors, there is limited research on melanoma.

**Table 2. Quantitative performance comparison across architectures (meta-analysis of reported results on ISIC)**

| Architecture | Avg AUC | Avg Sensitivity | Avg Specificity | Dataset(s) | Real-world validation? |
|---|---|---|---|---|---|
| VGG16 | 0.88 | 84% | 85% | ISIC 2016 | No |
| ResNet50 | 0.92 | 89% | 88% | ISIC 2017 | Limited |
| Inception-v3 | 0.91 | 87% | 89% | HAM10000 | No |
| Vision Transformer | 0.94 | 91% | 90% | ISIC 2019 | No |
| Ensemble (CNN+ViT) | 0.95 | 92% | 91% | Combined | No |

## 4. Challenges & Future

One of the problems that affects melanoma detection through modeling is the limited size and diversity, leading to class imbalance and poor representation of rare types of melanomas and skin colors [13], [28]. Datasets often suffer from insufficient numbers of labeled images because complex models require large quantities of information in order to be trained, leading to overfitting, which means models work better with training data rather than unseen cases, hampering their generalization in practice [37]. The issues of model generalization and consistency are complicated by the variety of imaging techniques and conditions. Overfitting due to limited amounts of diverse data and generalization of models among various demographics, imaging techniques, and skin conditions leads to poor diagnostic accuracy as one of the reasons behind the drawbacks of melanoma detection models [38], [39]. As interpretability provides an opportunity to test and validate predictions based on medical rationale, it builds confidence among dermatologists while providing valuable insights into decision-making mechanisms [39], [40]. By solving concerns about "black-box" models and ensuring collaboration between AI and healthcare professionals for the ethical treatment of patients, interpretable models assist doctors in understanding, validation, and acceptance of diagnostics [39], [40].

Through inclusion in different demographics as well as different geographies, collecting and merging large sets of images that include many different skin shades, many types of melanomas, and different imaging scenarios will contribute to creating databases that can help improve diversity. This will lead to less bias and greater generalizability of the model to all kinds of patients. Data sharing collaborations, unification of many datasets, proper annotation, and incorporating information from other modalities, including metadata, will contribute to creating more inclusive databases. [39]

There exist immense opportunities to enhance melanoma detection through novel approaches such as XAI techniques and federated learning methods. The XAI techniques, including Grad-CAM, LIME, and SHAP, which offer visible explanations for the workings of the machine learning model, increase the trust of clinicians and make validation possible, an important step towards implementation in practice. [41], [42]. Through federated learning, cooperation in training models using different data sets without sharing actual patient data not only addresses data security and legal concerns but also increases the pool of training data. [42].

The effective use of artificial intelligence models in melanoma detection calls for excellent performance validated across different populations and imaging settings, as well as interpretability, which serves as a link between AI results and clinical decisions. Some key issues are ongoing monitoring of bias and drift in models, integration into existing clinical practices, education for clinicians, and regulatory considerations. Furthermore, friendly user interfaces and explanation facilities may improve physician and patient safety and, in the long run, benefit patients through better accuracy. [42], [43].

In contrast to technical challenges, examples of workflow challenges include non-standardized imaging techniques, no integration of AI results with clinical reporting systems, and unclear legal responsibility for misinterpreting AI results.

### *4.1. Consent and Privacy of Patient Information*

In the context of data privacy, consent for the use of patient information in AI models training is often not provided. It is required by regulatory bodies (GDPR, HIPAA), but high-resolution dermoscopy images pose a threat of re-identification.

### *4.2. Regulatory Adoption*

To date, there is no approval from the FDA or CE Mark for any standalone diagnostic tool based on deep learning models for primary melanoma detection. Clinicians must monitor the systems.

### *4.3. Liability and Discrimination*

There is no clear liability when the AI algorithm fails to recognize a melanoma. Questions arise regarding the fairness of such approaches due to the bias of training datasets and low sensitivity to patients with non-white skin. Methods of explainable artificial intelligence for melanoma detection include:

- Grad-CAM – highlights areas relevant to classification;
- LIME – provides a local approximation of decision boundaries;
- SHAP – provides feature importance score.

The main limitation lies in the fact that the heat maps can be misleading if the model relies on artifacts rather than biological features, such as hair or ruler marks.

## 5. Conclusion

Given that models achieve an accuracy rate greater than 97% in commonly used benchmarks and match or surpass the level of dermatologist performance in a controlled setting, this research illustrates the tremendous progress that deep learning models have made during the past decade in detecting melanoma. Although the accuracy achieved by VGG, ResNet, and Inception models is considered the benchmark, vision transformers and ensemble learning may provide alternative directions toward incorporating global contextual information. It is worth noting that the contribution of the ISIC, HAM10000, and PH2 datasets, which have enabled a common benchmark, to the development of these models is undeniable. Several obstacles should be overcome to ensure that these models can be implemented for therapeutic purposes. Not only does the current dataset contain issues such as class imbalance and underrepresentation of different skin types, but also some rare subtypes of melanoma are missing. Model overfitting and poor results when applying them to other external datasets highlight the difference between benchmark performance and clinical application. Moreover, the "black box" nature of these models prevents them from being applied in practice because clinicians cannot verify the results generated by these models. The three main factors required for future success include creating globally representative and diverse data sets, incorporating XAI at the beginning stages of model creation, and constructing a framework of validation through prospective clinical trials. Cooperation among computer scientists, dermatologists, and regulatory authorities will be required to create melanoma diagnosis systems.

With federated learning, issues of privacy and legalities associated with patient data will no longer present an issue since training will not require sharing any patient information. Continuous learning enables model improvement without losing past knowledge through the continual addition of new dermoscopic instances. While both have yet to prove their value for melanoma detection, they are crucial components.

## References

[1] Colton Connor et al., “Clinical Approaches for the Management of Skin Cancer: A Review of Current Progress in Diagnosis, Treatment, and Prognosis for Patients with Melanoma,” *Cancers*, vol. 17, no. 4, p. 707, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[2] Layla Pires et al., “Time-resolved Fluorescence Lifetime for Cutaneous Melanoma Detection,” *Biomedical Optics Express*, vol. 5, no. 9, pp. 3080-3089, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[3] Hiam Alquran et al., “The Melanoma Skin Cancer Detection and Classification using Support Vector Machine,” *2017 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies*, Aqaba, Jordan, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[4] Nikki R. Adler et al., “Methods of Melanoma Detection and of Skin Monitoring for Individuals at High Risk of Melanoma: New Australian Clinical Practice,” *The Medical Journal of Australia*, vol. 210, no. 1, pp. 41–47, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[5] Abdul Sajid Mohammed et al., “The Smart Deep Learning based Model for Early Detection and Diagnosis of Melanoma,” *2024 International Conference on Integrated Circuits and Communication Systems*, Raichur, India, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[6] Clio Dessinioti et al., “Association of Skin Examination Behaviors and Thinner Nodular vs Superficial Spreading Melanoma at Diagnosis,” *JAMA Dermatol*, vol. 154, no. 5, pp. 544-553, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[7] JennaE. Rayner et al., “Clinical Perspective of 3D Total Body Photography for Early Detection and Screening of Melanoma,” *Frontiers in Medicine*, vol. 5, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[8] Lei Bi et al., “Automatic Melanoma Detection Via Multi-scale Lesion-biased Representation and Joint Reverse Classification,” *2016 IEEE 13th International Symposium on Biomedical Imaging*, Prague, Czech Republic, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[9] Scott W. Menzies, “Cutaneous Melanoma: Making a Clinical Diagnosis, Present and Future,” *Dermatologic Therapy*, 2006. [CrossRef] [Google Scholar] [Publisher Link]

[10] Serena Federico et al., "Advancements in Diagnosis of Neoplastic and Inflammatory Skin Diseases: Old and Emerging Approaches," *Diagnostics*, vol. 15, no. 16, p. 2100, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[11] Haley D. Heibel, Leah Hooey, and Clay J. Cockerell, "A Review of Noninvasive Techniques for Skin Cancer Detection in Dermatology," *American Journal of Clinical Dermatology*, vol. 21, pp. 513–524, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[12] R. Deepa et al., "Early Detection of Skin Cancer Using AI: Deciphering Dermatology Images for Melanoma Detection," *AIP Advances*, vol. 14, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[13] Hoda Naseri, and Ali A. Safaei, "Diagnosis and Prognosis of Melanoma from Dermoscopy Images using Machine Learning and Deep Learning: A Systematic Literature Review," *BMC Cancer*, vol. 25, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[14] Adekanmi A. Adegun, and Serestina Viriri, "Deep Learning-Based System for Automatic Melanoma Detection," *IEEE Access*, vol. 8, pp. 7160–7172, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[15] Tasneem Alkarakatly et al., "Skin Lesions Identification Using Deep Convolutional Neural Network," *2019 International Conference on Advances in the Emerging Computing Technologies*, AI Madinah AI Munawwarah, Saudi Arabia, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[16] Karl Thurnhofer-Hemsi, and Enrique Domínguez, "A Convolutional Neural Network Framework for Accurate Skin Cancer Detection," *Neural Processing Lettters*, vol. 53, pp. 3073-3093, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[17] Tri Cong Pham et al., "A Comparative Study for Classification of Skin Cancer," *2019 International Conference on System Science and Engineering*, Dong Hoi, Vietnam, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[18] Tallha Akram et al., "A Multilevel Features Selection Framework for Skin Lesion Classification," *Human-Centric Computing and Information Sciences*, vol. 10, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[19] Saleh Albahli, "A Robust YOLOv8-Based Framework for Real-Time Melanoma Detection and Segmentation with Multi-Dataset Training," *Diagnostics*, vol. 15, no. 6, p. 691, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[20] Alireza Golkarieh et al., "Semi-supervised GAN with Hybrid Regularization and Evolutionary Hyperparameter Tuning for Accurate Melanoma Detection," *Scientific Reports*, vol. 15, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[21] Abdulmateen Adebiyi et al., "Accurate Skin Lesion Classification Using Multimodal Learning on the HAM10000 and ISIC 2017 Datasets," *MedRxiv*, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[22] Muhammad Sajid et al., "Enhancing Melanoma Diagnostic: Harnessing the Synergy of AI and CNNs for Groundbreaking Advances in Early Melanoma Detection and Treatment Strategies," *International Journal of Imaging Systems and Technology*, vol. 35, no. 1, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[23] Javed Rashid et al., "Skin Cancer Disease Detection Using Transfer Learning Technique," *Applied Sciences*, vol. 12, no. 11, p. 5714, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[24] Olusola Oluwakemi Abayomi-Alli et al., "Malignant Skin Melanoma Detection using Image Augmentation by Oversampling in Nonlinear Lower-dimensional Embedding Manifold," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 29, no. SI-1, pp. 2600–2614, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[25] Khadija Nawaz et al., "Skin Cancer Detection using Dermoscopic Images with Convolutional Neural Network," *Scientific Reports*, vol. 15, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[26] Shivani Tyagi, and Sangeeta Mittal, "Sampling Approaches for Imbalanced Data Classification Problem in Machine Learning," *Proceedings of ICRIC 2019*, pp. 209-221, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[27] Pattaramon Vuttipittayamongkol et al., "Overlap-Based Undersampling for Improving Imbalanced Data Classification," *Intelligent Data Engineering and Automated Learning-IDEAL 2018*, pp. 689-697, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[28] Chun-Te Lu et al., "Skin Cancer: Epidemiology, Screening and Clinical Features of Acral Lentiginous Melanoma (ALM), Melanoma in Situ (MIS), Nodular Melanoma (NM) and Superficial Spreading Melanoma (SSM)," *Journal of Cancer*, vol. 16, no. 13, pp. 3972-3990, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[29] Niha Kamal Basha et al., "Mask Region-based Convolutional Neural Network and VGG-16 Inspired Brain Tumor Segmentation," *Scientific Reports*, vol. 14, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[30] Ahmad Naeem et al., "SCDNet: A Deep Learning-Based Framework for the Multiclassification of Skin Cancer Using Dermoscopy Images," *Sensors*, vol. 22, no. 15, p. 5652, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[31] Arief Budhiman, Suyanto Suyanto, and Anditya Arifianto, "Melanoma Cancer Classification Using ResNet with Data Augmentation," *2019 International Seminar on Research of Information Technology and Intelligent Systems*, Yogyakarta, Indonesia, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[32] Ananda Ananda et al., "Classification and Visualisation of Normal and Abnormal Radiographs; A Comparison between Eleven Convolutional Neural Network Architectures," *Sensors*, vol. 21, no. 16, p. 5381, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[33] Gelan Ayana, and Se-Woon Choe, "BUViTNet: Breast Ultrasound Detection via Vision Transformers," *Diagnostics*, vol. 12, no. 11, p. 2654, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[34] Satoshi Takahashi et al., "Comparison of Vision Transformers and Convolutional Neural Networks in Medical Image Analysis: A Systematic Review," *Journal of Medical Systems*, vol. 48, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[35] Seyed Mohammad Alizadeh, and Ali Mahloojifar, "Automatic Skin Cancer Detection in Dermoscopy Images by Combining Convolutional Neural Networks and Texture Features," *International Journal of Imaging Systems and Technology*, vol. 31, no. 2, pp. 695-707, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[36] Erdem Yanar, Firat Hardalaç, and Kubilay Ayturan, "CELM: An Ensemble Deep Learning Model for Early Cardiomegaly Diagnosis in Chest Radiography," *Diagnostics*, vol. 15, no. 13, p. 1602, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[37] Md Abdullah All Mahmud et al., "Explainable Deep Learning Approaches for High Precision Early Melanoma Detection using Dermoscopic Images," *Scientific Reports*, vol. 15, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[38] An Thi Phuong Nguyen, Rasel Mahmud Jewel, and Arjina Akter, "Comparative Analysis of Machine Learning Models for Automated Skin Cancer Detection: Advancements in Diagnostic Accuracy and AI Integration," *The American Journal of Medical Sciences and Pharmaceutical Research*, vol. 7, no. 1, pp. 15–26, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[39] Geetika Munjal et al., "SkinSage XAI: An Explainable Deep Learning Solution for Skin Lesion Diagnosis," *Health Care Science*, vol. 3, no. 6, pp. 438-455, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[40] Sutong Wang et al., "Interpretability-Based Multimodal Convolutional Neural Networks for Skin Lesion Diagnosis," *IEEE Transactions on Cybernetics*, vol. 52, no. 12, pp. 12623–12637, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[41] Syed Adil Hussain Shah et al., "Explainable AI-Based Skin Cancer Detection Using CNN, Particle Swarm Optimization and Machine Learning," *Journal of Imaging*, vol. 10, no. 12, p. 332, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[42] Eleni Briola et al., "A Federated Explainable AI Model for Breast Cancer Classification," *Proceedings of the 2024 European Interdisciplinary Cybersecurity Conference*, pp. 194-201, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[43] Shaik Balkhis Banu, and R. Murugesan, "Explainable Artificial Intelligence for Clinical Decision Support in Cardiac Rhythm Device Programming," *Rademics Research Institute*, 2025. [CrossRef] [Publisher Link]