



Artificial intelligence in the detection of skin cancer: State of the art

Michał Strzelecki, PhD^{a,*}, Marcin Kociołek, PhD^a, Maria Strąkowska, PhD^a,
Michał Kozłowski, PhD^b, Andrzej Grzybowski, MD, PhD^c, Piotr M. Szczypiński, PhD^a

^a Institute of Electronics, Lodz University of Technology, Łódź, Poland

^b Department of Mechatronics and Technical and IT Education, Faculty of Technical Science, University of Warmia and Mazury, Olsztyn, Poland

^c Institute for Research in Ophthalmology, Foundation for Ophthalmology Development, Poznań, Poland

Abstract The incidence of melanoma is increasing rapidly. This cancer has a good prognosis if detected early. For this reason, various systems of skin lesion image analysis, which support imaging diagnostics of this neoplasm, are developing very dynamically. To detect and recognize neoplastic lesions, such systems use various artificial intelligence (AI) algorithms. This area of computer science applications has recently undergone dynamic development, abounding in several solutions that are effective tools supporting diagnosticians in many medical specialties. In this contribution, a number of applications of different classes of AI algorithms for the detection of this skin melanoma are presented and evaluated. Both classic systems based on the analysis of dermatoscopic images as well as total body systems, enabling the analysis of the patient's whole body to detect moles and pathologic changes, are discussed. These increasingly popular applications that allow the analysis of lesion images using smartphones are also described. The quantitative evaluation of the discussed systems with particular emphasis on the method of validation of the implemented algorithms is presented. The advantages and limitations of AI in the analysis of lesion images are also discussed, and problems requiring a solution for more effective use of AI in dermatology are identified.

© 2024 Elsevier Inc. All rights reserved.

Introduction

Skin cancer ranks among the most common types of cancer, and it is recognized for its highly malignant nature. Its prevalence has outpaced that of nearly all other cancers, with annual rates rising by approximately 3% to 7% in fair-skinned populations over the past few decades.^{1,2} Consequently, there is a pressing need to conduct extensive screening tests to detect skin cancer at early stages. To facilitate this, it is crucial to automate the diagnosis of skin moles through

the utilization of image analysis and computational artificial intelligence (AI) techniques.

Early detection of melanoma, the most aggressive type of skin cancer, has a significant impact on patient prognosis. It increases the chances of successful treatment, prevents disease progression, and reduces the risk of life-threatening complications. Early diagnosis allows for less invasive surgical procedures, shorter recovery time, and effective adjuvant therapies. Regular skin examinations and prompt intervention are crucial for improving patient outcomes.³

Artificial intelligence has emerged as a transformative force in various fields of human activity, including medical diagnostics. In dermatology, the application of AI techniques has garnered significant attention, leading to progress in automated diagnosis and management of skin conditions.

* Corresponding author.

E-mail address: michal.strzelecki@p.lodz.pl (M. Strzelecki).

Accurate and timely detection of skin lesions, including melanoma and other types of skin cancer, is crucial for early intervention and improved patient outcomes.

Two distinct approaches in AI support skin lesion diagnostics. The first one involves machine learning,⁴ including neural networks and, especially, various types of convolutional neural networks (CNNs),⁵ which require numerous and reliable learning data.^{4,5} These methods, however, often create black-box systems where reasoning is not understandable for human experts. The second approach uses expert systems that rely on information ontologies.⁶ These systems employ understandable rules and dependencies, formalizing expert knowledge. Although expert systems offer transparency and interpretability, their preparation is labor-intensive.

The integration of AI techniques, such as machine learning and expert systems, with dermatology has led to advancements in automated analysis of skin lesion images. These AI algorithms have demonstrated promising results in various tasks, including skin lesion classification, segmentation, feature extraction, and risk prediction. By leveraging large data sets of annotated skin lesion images, AI models are capable of learning complex patterns and discriminating between benign and malignant cases. The ability to extract relevant features from skin lesion images enables AI systems to provide valuable insights into lesion characteristics, aiding dermatologists in their own decision-making process.

Methodology

This contribution aims to present novel perspectives on skin lesion analysis using AI. Alongside established methods, we will explore emerging trends and innovative techniques. These new perspectives may encompass various aspects, such as the integration of multimodal data, the use of deep learning architectures, and the incorporation of advanced image-processing algorithms. By presenting them, we aim to inspire further research and to foster the development of more effective and reliable AI-based solutions for dermatologic applications.

In the section on modern classification we have shown a few examples of recent work on the classification of skin lesions in dermatoscopic images. We have limited our brief review to documents published since 2019 that deal with the classification of melanocytic lesions. Additional limitations in the choice were the description of the method used in a way that allowed the experiment to be repeated and the research to be carried out for a wide set of data.

Large data sets with annotated images of skin lesions

To develop methods of image analysis based on AI, especially in the case of methods using deep convolutional networks, it is necessary to have significantly numerous (at least

several thousand) sets of annotated training and test images. Although skin lesions are a common problem, individual scientific units usually have limited sets of annotated images of such lesions. Fortunately, there are several publicly available databases of annotated images with skin pigmentation lesions.

One of the most popular large data sets used for training and testing methods for skin lesion classification is HAM10000.^{7,8} This data set contains dermatoscopic images from different populations, acquired and stored by different modalities. The data set consists of 10,015 images, which can serve as a training set for academic machine learning purposes. Cases include a representative collection of all important diagnostic categories in the realm of pigmented lesions, including actinic keratoses and intraepithelial carcinoma/Bowen disease, basal cell carcinoma, benign keratosis-like lesions (solar lentigines/seborrheic keratoses and lichen-planus like keratoses), dermatofibroma, melanoma, melanocytic nevi, and vascular lesions (angiomas, angiokeratomas, pyogenic granulomas, and hemorrhage).

The images of the HAM10000 set were collected over a period of 20 years from two different sites: the Department of Dermatology at the Medical University of Vienna, Austria, and the skin cancer practice of Cliff Rosendahl in Queensland, Australia. Some of the images originate from the time before the era of digital cameras, so they were stored as diapositives and later converted to digital form. All images are stored as JPEG files with sizes of 800 × 600 pixels. More than 50% of lesions have been confirmed by pathology; the diagnostic basis for the remaining cases was either followup, expert consensus, or confirmation by *in vivo* confocal microscopy.

The other popular data sets come from a series of challenges hosted by the International Skin Imaging Collaboration (ISIC).⁹ The ISIC is an international effort to improve melanoma diagnosis, sponsored by the International Society for Digital Imaging of the Skin.¹⁰ Between 2016 and 2020, there were five such challenges, and each of them had its own set of data. Because this contribution concerns disease recognition, we will limit ourselves to data sets for disease classification, as in some challenges, tasks other than classification were considered.

The first challenge was held in 2016.¹¹ The full name of the challenge is “Skin Lesion Analysis Toward Melanoma Detection: A Challenge at the International Symposium on Biomedical Imaging (ISBI).” Its data set is referred to as ISCS2016 or ISBI2016. Lesion classification data included the original image accompanied by the definitive malignancy diagnosis and the lesion segmentation mask. This had three parts: lesion segmentation, detection and localization of visual dermoscopic features/patterns, and disease classification. The data set for the challenge has grown over time. Details of the data set for the classification part are given in Table 1.

Figure 1 shows examples of images from the ISIC 2019 challenge.¹⁴

Table 1 Details of International Skin Imaging Collaboration challenge data sets

| Year | No. images, training | No. images, validation | No. images, testing | No. classes | Percentage of melanoma class | Data format | Image size |
|-------------------------|----------------------|------------------------|---------------------|-------------|------------------------------|--------------|----------------------------|
| 2016 ¹¹ | 900 | | 379 | 2 | 19.2% | JPEG | 576 × 786 to 4,288 × 2,848 |
| 2017 ¹² | 2,000 | 150 | 600 | 2 | 18.7% | JPEG | 576 × 786 to 6,748 × 4,499 |
| 2018 ^{8,13} | 10,015 | 193 | 1,512 | 7 | 11.1% | JPEG | 640 × 450 |
| 2019 ^{8,12,14} | 25,331 | | 8,238 | 9 | 17.9% | JPEG | 576 × 786 to 1,024 × 1,024 |
| 2020 ^{15,16} | 33,126 | | 10,982 | 9 | 1.8% | JPEG + DICOM | 640 × 480 to 6,000 × 4,000 |

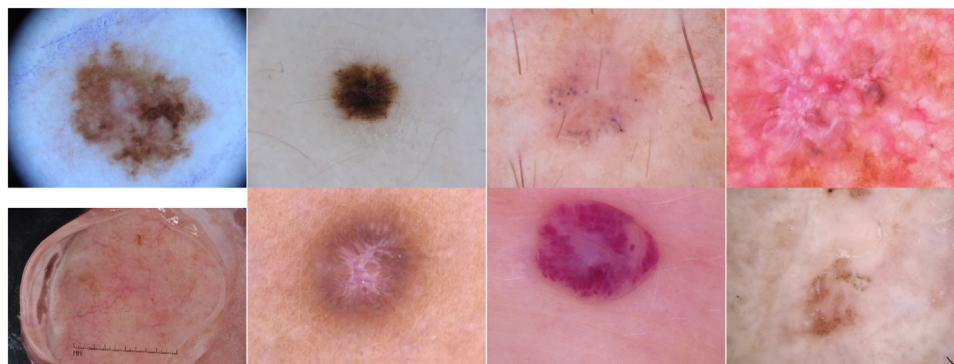


Fig. 1 Examples of images from the ISIC 2019 challenge.¹⁴ The upper row from the left: melanoma, melanocytic nevus, basal cell carcinoma, actinic keratoses. The bottom row from the left: benign keratosis-like lesions, dermatofibroma, vascular lesions, squamous cell carcinoma. ISIC, International Skin Imaging Collaboration.

Some less popular data sets are hosted by the DermNet service from New Zealand.¹⁷ DermNet describes itself as “the world’s leading free dermatology resource . . . created for healthcare professionals to serve as a comprehensive database of skin conditions.” This data set contains slightly more than 20,000 images obtained by various techniques; however, there are only 191 images of melanoma. Free images are watermarked, and clean images can be downloaded for a fee. Dermofit¹⁸ is a dermoscopic image library provided by the University of Edinburgh. It consists of 1,300 skin images containing 10 types of skin lesions, including 331 images of melanoma. These images can be downloaded under an academic license for a fee.

Ontology-based expert systems

Researchers have increasingly focused their efforts on developing formal computational frameworks that express their expertise, knowledge, concepts, and relationships in a manner that is readable and applicable to decision-making expert systems in the medical field. In the domain of skin-related diseases, digital skin images exhibit diverse semantic features, presenting a considerable challenge in understanding

symptoms and detecting early-stage skin cancers. To address this, the formalization and rendering of these standards in a machine-readable format through the use of ontologies becomes imperative.

Ontologies serve as formal representations that define categories, properties, and relationships between concepts within specific subject areas. They play a crucial role in providing a precise and shared vocabulary for the semantic annotation of resources and describing them through metadata. In dermatology, biomedical ontologies provide essential domain knowledge that facilitates data integration, information retrieval, data annotation, natural language processing, and decision support. Considering the visual nature of dermatology and its substantial reliance on visual analysis, the annotation of images plays a crucial role in medical decision expert systems. This underscores the value of formalizing ontologies to build a knowledge base that facilitates inference and decision-making processes.

The ABCD rule forms the foundation for dermatologists’ diagnoses, encompassing the Asymmetry, Border structure, variegated Color, and Differential structure characteristics of skin lesions. This rule is generally formulated, however, and diagnosing skin lesions requires a deeper understanding of subtle differences, which dermatologists have acquired

through experience. Creating a strict, detailed, and precise description that enables accurate differentiation and classification of various skin lesions demands substantial work and the formalism provided by ontologies.

Biomedical ontologies play a vital role in providing essential domain knowledge for driving data integration, information retrieval, data annotation, natural language processing, and decision support. To enable inference over classes, description logic (DL) attributes, domain, and range constraints are used. The ontology for skin lesions¹⁹ was implemented in OWL format, leveraging the capabilities of DL to establish a standardized syntax that eliminates semantic ambiguity and is suitable for a robust inference engine. A DL-based approach offers distinct advantages, especially in the interpretation of medical images, where extensive domain knowledge is involved.

The proposed ontology model adopts a multilevel hierarchy, where a collection of graphical descriptors encompasses a set of image subjects, which, in turn, encompasses a set of medical terminology. The authors specifically focus on image attributes related to the ABCD principle, such as contour features, asymmetry, surface color, and texture features. The ontology serves as the foundation for developing an expert system that supports the annotation of melanocytic changes. The implemented program calculates image attributes and performs inference by considering the semantic relationships between the quantitative attributes of the images (data) and the quality (classes) of the changes.

The presented results¹⁹ demonstrate the precision and recall values obtained when distinguishing cases of melanoma, dysplastic changes, and common skin nevi. For melanoma, a precision of 0.75 and a reproducibility of approximately 0.65 were achieved. In the case of dysplastic changes, the precision exceeded 80%, and the reproducibility was more than 0.70. For common changes, the precision reached 0.80, with a reproducibility of more than 0.65.

The authors of the contribution²⁰ introduce an ontology that specifically focuses on the classification of skin nevi and its semantic annotation. Drawing inspiration from dermatology experts, the Bag-of-Words (BoW) technique models these concepts in diagnosing skin lesions. The BoWs are constructed based on the features extracted from the lesion images, encompassing shape, color, and texture descriptors. A machine learning support vector classifier generates the BoWs from these features. The authors also derived their ontology from the scoring system of the ABCD rule. Experiments with the developed annotation system were conducted on a publicly available database comprising 206 lesion images, demonstrating that the ontology provided an efficient framework for analysis. Incorporating semantic relations between concepts facilitated the integration of expert knowledge, leading to a more appropriate approach for classifying lesion severity. A comparison was made between the results obtained using the proposed method and those achieved by other approaches. The presented approach demonstrated a high level of accuracy equal to 0.77% and a sensitivity of

0.97; however, the obtained specificity of 0.48 can be regarded as unsatisfactory.

The issue of insufficient specificity was addressed by implementing fuzzy reasoning principles.^{21,22} In this regard, the developed ontology was enhanced by incorporating vague decision aspects. The recognition task, differentiating between the melanoma and benign classes, was performed using a k -nearest neighbor classifier with a parameter k set to 7. This approach yielded an accuracy of 0.92 and a sensitivity of 0.96. The specificity increased to 0.89, which is a promising outcome. To evaluate the performance of this approach, a comparison was made with results obtained from other algorithms, including two convolutional neural networks. The findings demonstrated that using the fuzzy decision approach and ontology-based inferencing is a viable alternative to these networks.

(Editor's note: The aforementioned systems may be enhanced by using the ABCDE rule, in which the "E" stands for "evolution." Applying this rule requires that the appearance be evaluated over time, preferably by taking good clinical and dermoscopic images at each observation. Changes in the appearance of the lesion, particularly expansion, are considered signs of malignancy.)

Modern classification systems for dermoscopic images of skin lesions

The Global-Part Convolutional Neural Network (GP-CNN)²³ model integrates fine-grained local and global context information. The Global-Part model comprises a Global Convolutional Neural Network (G-CNN) and a Part Convolutional Neural Network (P-CNN). The G-CNN interacts with downscaled dermoscopy images to extract information at the global scale and generate the Classification Activation Map (CAM). The P-CNN is trained with CAM-guided image patches to capture local-scale information in skin lesion regions. An ensemble strategy is employed to enhance the classification performance further. The ensemble combines discriminant information from GP-CNNs separately trained on the same image set.

In another approach, the Geometric Active Contour (GAC)²⁴ was used to segment the area of lesions and isolate them from healthy skin. Subsequently, pretrained standard CNN models, AlexNet, GoogleNet, and VGG16, followed by global average pooling, are applied to generate features. Feature vectors are then reduced through principal component analysis (PCA) and, finally, fed into the artificial neural network (ANN) and random forest (RF) classifiers.

The fusion of three deep learning architectures²⁵—Xception, Inception-ResNet-V2, and NasNetLarge—pretrained on images from ImageNet is employed for the Large-Scale Visual Recognition Challenge 2012.²⁶ Subsequently, transfer learning is applied to all these architectures using data from the ISIC 2019 database. All

layers except the last ones undergo transfer learning, while the output layer is trained from scratch. In the next step, a LightGBM framework, a version of Gradient Boosted Decision Tree (GBDT), is trained using the features extracted from the last pooling layer of each deep learning model. As in the previous step, score-level averaging is applied to combine the prediction scores assigned to each class for the augmented images from the test set, locally, within a single network, and globally among different models.

Another approach uses a multistep algorithm.²⁷ The first step involves preprocessing, including contrast enhancement through fast local Laplacian filtering, followed by a transformation to the HSV (Hue, Saturation, Value) color space, which is performed. In the second step, segmentation into lesion and healthy skin areas is performed using a transfer-learned VGG19 network operating on the images from the first step. Subsequently, the original image is masked by the segmentation outcome so that healthy skin pixels from the original image are set to zero. Finally, the Inception V3 network is employed to obtain the feature vector. These feature vectors are then reduced based on entropy, and a fully connected network is applied for classification into three classes.

Three popular CNNs, DenseNet-201, Inception-ResNet-V2, and Inception-V3,²⁸ were trained on the ImageNet data set.²⁶ The authors tested seven different optimizers and four different numbers of dense layers. Two training sets were used: one with the original HAM10000 data set and the second with an augmented HAM10000 data set in which the numbers of samples in classes were balanced. Two training approaches were tested. In the first, only dense layers were trained from scratch, whereas in the second approach, selected convolutional layers were transfer-learned, and the dense layers were trained from scratch. The best results were reported for DenseNet 201 with transfer learning on an augmented balanced data set.

A modified GoogleNet²⁹ was used to classify images of skin lesions.³⁰ GoogleNet was pretrained on the ImageNet²⁶ data set. The modification of the GoogleNet architecture consisted of replacing the last three layers with a new fully connected layer with 2,048 neurons, followed by the output classification layer with the SoftMax activation function. This new network was transfer-trained on the ISIC 2019 data set. Improvement in the balance of training data was obtained in two ways: in the first, the minor numerous classes were augmented; in the second, the data for individual classes were randomly reduced, adjusting their number to the lowest numbered class.

The DermoExpert³¹ is a two-stage framework. The first step is preprocessing, which involves segmentation, rebalancing, and augmentation. The segmentation uses DSNet.³² A weighted loss function is used for rebalancing, and augmentation relies on rotation and intensity changes. After the preprocessing, a so-called Hybrid CNN classifier is employed. This classifier comprises three custom-designed parallel CNN networks called feature generators. The outputs of the feature generators are fused through channel concate-

nation. The fused output and the separate outputs of the feature generator are processed by global average pooling layers and provided to four identical but separate dense networks with three fully connected layers, followed by the output layer with a SoftMax activation function. The outputs of these thick layers are averaged and provided as the final output. The weights of the feature generators were pretrained on ImageNet data.²⁶

The best algorithms for recognizing skin lesions can be found at the top of the leaderboard from ICIS challenges, especially the latest one, SIIM-ISIC Melanoma Classification from 2020.¹⁰ Unfortunately, the challenge with these algorithms is that they are rarely documented in scientific contributions. Occasionally, summaries or reports are available to describe the approach taken. On the positive side, the training data set (images plus classification) is extensive, publicly available, and thoroughly described on the challenge website. The test set was provided without classification, and the results were submitted to the challenge organizers for evaluation, ensuring objectivity. Many participants adopt an approach to publish the complete code used during the challenge and their report, allowing others to replicate the entire learning and testing process. The majority of participants at the top of the results ladder employ an ensemble of classifiers, with most or all of the classifiers being deep convolutional networks. This trend is evident even in the case of the winner³³ of the ISIC 2021 challenge, whose code is accessible on the open GitHub repository.³⁴

There are also some problematic works; for example, in one contribution,³⁵ the authors used the HAM10000 data set. The images were resized to 28×28 pixels, reducing some lesions to 20 pixels. To balance the classes, all classes except melanocytic nevi were augmented by flipping and random rotation in the range of $\pm 10^\circ$. This expanded data set was divided into 2/3 for the training set, 1/6 for the validation set, and 1/6 for the test set. A relatively simple CNN with five convolutional layers, one fully connected layer, and an output layer was used for classification. A max-pool layer followed each convolutional layer. The entire network had about 172,000 parameters. Reported results are excellent, with an average accuracy, precision, recall, and F1 score of 0.98.

A closer look at the classification parameters of individual classes indicated that the best results were achieved for the most heavily augmented classes. For example, for the class Dermatofibroma, which was supplemented more than 50 times, all reported classification measures were 1. The least increased class was basal cell carcinoma, supplemented more than seven times, and reported classification measures were between 0.94 and 1. In the case of the melanocytic nevi class, accuracy and reproducibility were at the level of 0.87, the F1 score was at the level of 0.93, and precision was at the level of 1. Considering this, after the proposed augmentation of images with such low resolution, they did not differ much from the original, and the network was tested on known images. Such a significant reduction in resolution

is a conceptual error because some structures necessary for humans to recognize melanocytic lesions have been blurred here.

An interesting study focused on the improvement in the diagnosis of skin lesions supported by analysis using CNNs.³⁶ In this research, the commercial FotoFinder³⁷ system was used. This system holds an MDR (Medical Device Risk) Class IIa certificate and CE (Conformité Européenne) mark, making it approved for medical use in the European Union (EU). With its assistance, 188 patients from two medical units involving 22 dermatologists in the study were examined over two years. A total of 228 suspicious melanocytic lesions were detected in the examined patients, of which 38 were melanoma.

The research demonstrated that diagnostic accuracy was significantly improved when CNN-based skin lesion analysis was incorporated into decision making. The mean sensitivity improved from 84.2% to 100%, the mean specificity improved from 72.1% to 83.7%, and the mean accuracy improved from 74.1% to 86.4%. Unfortunately, the authors did not provide details on the CNN network used, and the data set is limited.

Table 2 summarizes publications on skin lesion classification methods. Accuracy measures presented here are specifically for the detection of malignant melanoma. Many existing publications use average measures to classify various moles, which can often inflate these measures due to unbalanced data sets. We chose to focus on measurements related to melanoma detection because it poses the greatest threat to patients, unlike other skin lesions.

The table excludes contributions in which we identified methodologic problems and those with inaccurate descriptions of the methods used, rendering the experiment impossible to reproduce. The last entry in the table pertains to the winner of the ISIC 2020 challenge. No scientific publication is available in this instance, but the authors have provided a brief report with the complete code accessible in an open repository.

Smartphone applications

Smartphone devices are widely used and accessible portable computers with enriched connectivity options. Equipped with cameras and more computational power, they are becoming increasingly helpful in various medical applications. Some of them are assisting in skin self-examination, tracking the evolution of suspicious lesions, or being used for storage and forward teledermatology. To increase the quality of the images, there are also specialized lens adapters—mobile dermatoscopes, such as DermLite³⁸ or MoleScope³⁹—that can zoom in on suspicious lesions. These devices enable patients to monitor their skin lesions and moles from home and send high-quality images for regular checkups. They allow one to take photos with high magnification and use cross-polarized or polarized

light. There are also dedicated applications and databases in which the doctor can store, compare, or analyze these images.⁴⁰⁻⁴²

The developed systems are often implemented as mobile applications to expand their availability and the ability to collect new data. Basic information on the most significant research is summarized in Table 3.

The first attempt to implement an automatic classification system in mobile applications used a traditional approach for the parametric description of skin lesions.⁴³ Iowa State University researchers developed an application that extracts the ABCD features (Asymmetry, Border irregularity, Color variegation, and Diameter) from photographed lesions. The smartphone's detachable 10 × lens is used to enhance the image quality. Before the feature extraction operation, image preprocessing and segmentation are performed. The extracted features are input data for the support vector machine (SVM) classifier. The authors tested three kernels for the classifier: linear, radial basis function (RBF), and polynomial against evaluation parameters. The training and testing of the system were performed using the publicly available database that consists of 200 dermatoscopic images (80 atypical nevi, 80 common nevi, and 40 melanomas). To balance the number of samples in the classes, the data augmentation using the Synthetic Minority Oversampling Technique (SMOTE) was performed. The best result of the system was achieved using an SVM classifier based on an RBF kernel: sensitivity of 0.8, specificity of 0.90, and accuracy of 0.88.

Another smartphone application⁴⁷ uses a convolutional neural network (CNN). The network architecture evaluation was performed on a total of 8,000 images. The CNN classifier obtained an accuracy of 0.788, a sensitivity of 0.913, and a specificity of 0.73. This part of the presented work seems to add functionality to the developed application rather than its core value. The emphasis of the work has been placed on the standard approach of image processing, ABCD feature extraction, and its presentation in the augmented reality application that works in real time on the smartphone device.

The mobile application can also act only as a front-end of the designed system.⁴⁸ The data processing and neural network calculations are performed on the remote server to which the smartphone client application connects. The primary role of the mobile app is to gather images of the skin lesions and provide simple clinical descriptions. After that, it sends them to the server to add them to the database or to get the result of processing and diagnosis. The clinical information added to each image of the lesion is its description with yes/no data referring to increased (over time) pain, changes, bleeding, and elevation of its surface. The PAD-UFES-20 data set with 658 cancers and 1,399 noncancer lesions, 2,057 in total, confirmed this. The CNN used was modified using ResNet50, as described in a contribution published in 2020.⁴⁹ The authors described in detail the approach for CNN training and testing, including CNN structure and parameters,

Table 2 Summary of specific classification parameters obtained for melanoma class

| Reference | Data set | No. classes used for training | AI algorithm | Precision | Accuracy | Sensitivity | Specificity | Area under the curve |
|---|-------------------------------|-------------------------------|---|----------------|----------|----------------|-------------|----------------------|
| Tang et al (2020) ²³ | ISIC 2016 ISIC 2017 | 3 | Own approach | - | 0.850 | 0.376 | 0.965 | 0.891 |
| Olayah et al (2023) ²⁴ | ISIC 2019 | 8 | Standard CNNs | 0.887 | 0.961 | 0.889 | 0.994 | 0.944 |
| Ahmed et al (2020) ²⁵ | ISIC 2019 | 9 | Ensemble of standard CNNs | 0.903 | - | - | - | 0.925 |
| Saba et al (2019) ²⁷ | PH2 ISIC 2016 ISIC 2017 | | VGG19 followed by Inception V3 | 0.950 | 0.948 | 0.945 | 0.980 | 0.980 |
| Villa-Pulgarin et al (2021) ²⁸ | HAM10000 ISIC 2019 | 7, 8 | DenseNet-201, Inception-ResNet-V2, and Inception-V3 | 0.960 | - | 0.940 | - | - |
| Kassem et al (2020) ³⁰ | ISIC2019 | 8 | Modified GoogleNet | 0.804 | 0.949 | 0.798 | 0.970 | - |
| Hasan et al (2022) ³¹ | ISIC 2016 ISIC 2017 | 2, 3 | Ensemble of CNNs | 0.820 0.730 | - | 0.610 0.620 | - | - |
| Ha et al (2020) ³³ | ISIC 2020 | 2 | Ensemble of CNNs | - | - | - | - | 0.949 |

AI, artificial intelligence; CNN, convolutional neural network; ISIC, International Skin Imaging Collaboration.

Table 3 Summary of methods used in a smartphone application for a skin examination

| Reference | Certification | AI algorithm | Data set | Performance | Classes |
|---|--------------------------|--|--|--|--|
| Kalva et al (2019) ⁴³ | - | SVM Kernels: linear, RBF, polynomial against evaluation parameters | PH2 database ⁴⁴ 200 dermoscopic images | Sensitivity = 0.80 Specificity = 0.90 Accuracy = 0.88 | Atypical nevi (80), common nevi (80), melanomas (40) |
| Udrea et al (2019) ⁴⁵ SkinVision ⁴⁶ | MDR Class IIa CE mark | Conditional adversarial network algorithm for segmentation SVM with RBF kernel for classification | 131,873 training, 285 cancers, 6,000 benign for validation | Sensitivity = 0.951 Specificity = 0.783 | Actinic keratosis, basal cell carcinoma, Bowen disease, malignant melanoma, squamous cell carcinoma |
| Francesse et al (2021) ⁴⁷ | - | CNN Four convolutional blocks + fully connected classification layer | 8,000 images | Sensitivity = 0.788 Specificity = 0.913 Accuracy = 0.73 | Actinic keratoses, Bowen disease, benign keratosis-like lesions, dermatofibroma, melanoma, melanocytic nevi, vascular lesions |
| Krohling et al (2021) ⁴⁸ | - | CNN ResNet50 with the last layer modified for skin cancer classification | PAD-UFES-20 ⁴⁹ 2,057 images | Accuracy = 0.85 Recall = 0.96 | Cancer: melanoma (67), basal cell carcinoma (442), squamous cell carcinoma (149) Noncancer: actinic keratosis (543), seborrheic keratosis (215), nevus (196), others (445) |
| Maduranga et al (2022) ⁵⁰ | - | MobileNet with transfer learning | HAM10000, ⁸ 6,705 images used | Accuracy = 0.85 | Dermatofibroma (115), vascular lesions (142), actinic keratoses (327), basal cell carcinoma (514), benign keratosis-like lesions (1,099), melanoma (1,113), melanocytic nevi (6,705) |
| Kränke et al (2023) ⁵¹ SkinScreener ⁵² | MDR Class IIa CE mark | Classic CNN and a novel stratification CNN based on a region proposal network | 2,057 images with metadata | Detect algorithm: Sensitivity = 0.964 Specificity = 0.949 Analyze algorithm: Sensitivity = 0.9534 Specificity = 0.903 | Benign (692), precancerous (183), malignant (196) |

AI, artificial intelligence; CE, Conformité Européene; CNN, convolutional neural network; MDR, medical device risk; RBF, radial basis function; SVM, support vector machine.

evaluation criteria, impact of clinical information, and data balancing.

The obtained results attained an accuracy of 0.85 and a reproducibility of 0.96. These values were the result of the approach where the clinical data were also taken into consideration as input data without the images. The improvement varies between evaluation parameters but falls between 1.4 and 2.4 pp.

Another approach used a pretrained CNN-type MobileNet network.⁵⁰ The fine-tuning training was performed using the HAM10000 database (10,000 images). It contains the images of seven classes of pigmented lesions (not only cancers), including dermatofibromas, vascular lesions, actinic keratoses, basal cell carcinomas, benign keratosis-like lesions, melanomas, and melanocytic nevi. Finally, more than 6,000 images were divided into training, validation, and testing subsets in the ratio of 60%/20%/20%, respectively. CNN was implemented and trained in a TensorFlow library with a Keras package. To implement the trained network into an Android application, it had to be converted to a TensorFlowLite library supported in an Android Studio IDE (Integrated Development Environment). It was claimed that the application achieved an accuracy of 0.85; however, the authors present neither a detailed specification of testing and accuracy calculation nor a training description. The CNN accuracy obtained with the use of smartphones was lower than with the use of computers. The authors concluded that this is a consequence of converting the trained model to be compatible with the lightweight version of the TensorFlow library.

One of the newer publications presents a user-friendly app for easy and quick cancer risk assessment of skin lesions.⁵¹ The patients under examination by the system were also assessed by two dermatologists or with the histology of the lesion, if available. These authors proposed a two-stage approach using a different neural network for each. The process was divided into image classifier (analysis) and image region proposal (detection). The analysis stage was based on CNN, while the detection of novel stratification CNN was developed on a region proposal network (RPN). The second one, more computationally demanding, takes around two to five seconds for processing on modern smartphones. A three-level decision for assessing images was introduced: subcategory, category, and risk level. Different vendors and models of smartphones were used for processing and acquisition of skin images. During this research, 238 patients with 1,171 lesions were examined, resulting in 18,384 images. The detected RPN model achieved a sensitivity and specificity of 96.4% and 94.85%, whereas CNN analysis yielded 0.954% and 0.903, respectively. The final application has been commercialized under SkinScreener; it has an MDR Class IIa medical device certification and complies with the strictest EU data protection requirements.

Last but not least is the SkinVision app, the most cited^{45,53,54} and reviewed smartphone application.^{42,43,55} SkinVision is a clinically validated, regulated medical device. It has an MDR Class IIa certificate, CE mark, TGA

(Therapeutic Goods Administration) approval, and ISO 13485:2016 certification. SkinVision sponsored a few clinical studies on the accuracy of the SkinVision app's risk assessment for University Hospital Munich, Catharina Hospital Eindhoven, clinical studies, and SkinVision app user databases.⁵⁴ This app was developed in 2011 and is still commercially available. The implemented algorithms were evaluated and improved for better sensitivity during this period. Since 2018, they have replaced the rule-based classification algorithm with a machine learning approach (Udrea et al, unpublished material, 2019). It uses a machine learning algorithm to analyze spots on the skin. A smartphone application's accuracy for triaging skin lesions based on machine learning algorithms has also been evaluated.⁴⁵ These results have demonstrated a sensitivity of 0.95 and a specificity of 0.78. More than 130,000 images were used for training. The Conditional Adversarial Network algorithms combined with an SVM classifier were used to segment the lesions.

Total body systems

Total body (or full-body, whole-body) systems (further referred to as TBS) represent an alternative for skin lesion analysis, including melanoma detection. Contrary to classic approaches, in which the dermatologist focuses on a given lesion characterized by a single image acquired by a dermatoscope, high-quality camera, or smartphone, TBS captures a series of photographs covering the patient's entire body. In this way, one can detect suspicious lesions and find lesions that have appeared since the last examination. Such followup examinations also make it possible to see lesions whose size has increased, which may suggest neoplastic changes. TBS are not commonly described in the literature, especially those that use AI-based algorithms.

One of the oldest systems was designed for lesion detection in optical images of larger body fragments⁵⁶ (arm, forearm). Image analysis has included skin detection, hair removal, and lesion detection based on difference-of-Gaussian filters, followed by the SVM classifier. The obtained results show sensitivity in the range of 0.80 to 0.85 and accuracy in the range of 0.76 to 0.79, depending on SVM working conditions.

A full-body system⁵⁷ was created to discriminate new and rapidly evolving melanocytic lesions and to implement a three-dimensional (3D) stereo capture system with 22 acquisition units. Further image preprocessing includes the removal of artifacts, particularly hair that might be visible in some photos. Next, Laplacian-over-Gaussian filtering was applied for lesion segmentation. The linear discriminant analysis was used for lesion classification—the system identifies “suspect” lesions that can be new and modified. The lesions were analyzed for different size groups (up to 3, 5, and 7 mm). The resulting recall for groups was 0.9, and precision varied from 0.5 for small new lesions to 0.8 for larger ones. The main limitations of this study were the use of artificial le-

sions for algorithm testing and system evaluation performed for a small number of patients (six men and six women).

Another TBS is designed to detect size change in multiple pigmented skin lesions using generated mole maps.⁵⁸ The system is based on a cabin equipped with 21 high-resolution cameras that capture the patient's body in 24 positions. The lesions are detected using the MSER (maximally stable extremal regions) algorithm. A set of complex algorithms, that is, lesion geometry, is applied for lesion matching and assessing its size change. The system performance was summarized by stating that the system detects almost 100% of the lesions and was tested on several patients. The system matching algorithm has subsequently improved by implementing rigid transformations of 3D point clouds and using non-rigid coordinate plane deformations in local ROIs (Region of Interest) around the lesions of interest.⁵⁹ Matching algorithms were tested in the melanoma unit of the Hospital Clínic de Barcelona Barcelona, Spain, on many thousands of lesions, reaching a recall of 0.82 and a precision of 0.999.

One contribution⁶⁰ presents a broad body image analysis to detect suspicious pigmented lesions. More extensive lesions (diameter > 3 mm) were manually segmented, and then ABCD features were calculated for each lesion in the image. Logistic regression was used for lesion classification, producing 0.84 sensitivity and 0.75 accuracy in detecting suspicious lesions. The system was tested on 133 patients and, altogether, 1,759 skin lesions.

A deep learning approach to detecting and classifying suspicious lesions in total body images was also implemented.⁶¹ The lesions were identified using a blob detection algorithm using Laplacian-of-Gaussians filtering and scale-invariant feature transformation. Different deep network architectures (CNN, VGG16, Xception with a transfer learning model) were implemented for identified lesion classifications. The six classes include nonsuspicious lesion types A and B (defined in the text), suspicious lesions, skin, skin edges, and background. The system was tested on many lesions (more than 33,000), reaching 0.80 average accuracy for lesions in the suspicious classification endpoint. The developed CNN algorithm also obtained good agreement with dermatologists' assessment.

Another prototype of the TBS system was developed by Skopia Esthetic Clinic, Kraków, Poland, in cooperation with the Institute of Electronics, Lodz University of Technology, Poland.⁶² To acquire the patient's pictures, the system is equipped with digital cameras that take 32 overlapping pictures that cover almost the whole area of the patient's skin, as shown in Figure 2. The prototype is equipped with image processing software that enables, among others, lesion detection and segmentation algorithms that use the Yolo deep network and an Otsu algorithm. The system also implements a procedure for building a 3D patient model. This permits acquired skin images with the model and orthorectify them for detecting size and shape changes in nevi.⁶³ It is also possible to see new lesions (or growing lesions) in followup studies. The lesion pairing algorithm implements the feature match-



Fig. 2 Example of device used for capturing pictures of the total area of human skin.⁶⁵

ing method (features are generated by scale-invariant feature transformation, BRISK, and AKAZE algorithms), and wrong matches are removed by the RANSAC algorithm.⁶⁴ The matching accuracy is 0.9, and the precision of lesion detection is 0.7 to 0.91, depending on lesion size. It is possible to detect changes in the size of moles of 1 mm.

In the literature, AI-based algorithms are also reported as convolutional neural networks⁶⁶ for analysis of the whole-body images to detect skin spots and classify them as nevi or non-nevi. The algorithm was tested on images from the Mind Your Moles study.⁶⁷ Images of 82 individuals (57,742 lesions) were considered in the training set, and data from 10 individuals (4,868 lesions) comprised the test set. The reported system performance of naevi detection (compared to senior dermatologists' assessment) shows sensitivity and specificity of 0.79 and 0.91, respectively, for lesions ≥ 2 mm and 0.84 and 0.91 for lesions ≥ 5 mm. Another such algorithm is described in another contribution.⁶⁸ It detects and longitudinally tracks skin lesions on 3D total-body skin surface scans from the 3DBodytex database. The algorithm was trained using Faster R-CNN. The obtained results show similar lesion detection accuracy when compared to three human annotators. The achieved accuracy for the matching algorithm was 0.88, and the accuracy of lesion detection was 0.71.

The performance of such systems is presented in Table 4. None of the discussed TBS is medically certified. The exception may be the ATBM master system (FotoFinder, Germany, <https://www.fotofinder.de/en/technology/total-body-dermoscopy/bodystudio-atbm/master>), which

Table 4 Performance of total body systems

| Reference | Data set | Lesion size (mm) | Lesion detection performance | Malignant lesion detection performance | Matching accuracy in followup studies | Three-dimensional patient model |
|--|---|------------------|--|--|---|---------------------------------|
| Taeg et al (2007) ⁵⁶ | 16 images | All | Sensitivity = 0.80-0.85 Accuracy = 0.76-0.79 | - | - | No |
| Bogo et al (2014) ⁵⁷ | 50 images of 12 patients | 1-3 >5 | Precision = 0.50 Sensitivity = 0.90 Precision = 0.80 Sensitivity = 0.90 | - - - | - - - | Yes |
| Korotkov et al (2019) ⁵⁹ | 240 overlapping images of 25 patients | >3 | - - | - | Sensitivity = 0.82 Precision = 0.999 | No |
| Birkenfeld et al (2020) ⁶⁰ | 3-8 images for 133 patients | All | - | Sensitivity = 0.84 Accuracy = 0.75 | - | No |
| Soenksen et al (2021) ⁶¹ | Single wide-field images of 133 patients | All | - | Accuracy = 0.97 | - | No |
| Strzelecki et al (2021) ⁶² | 480 images of 15 patients | 1-2 2-5 >5 | Precision = 0.91 Sensitivity = 0.05 Precision = 0.88 Sensitivity = 0.92 Precision = 0.70 Sensitivity = 0.68 | - | Sensitivity = 0.90 F1 score = 0.9 | Yes |
| Betz-Stablein et al (2022) ⁶⁶ | Images of 92 individuals from Mind Your Mole study ⁶⁷ | >2 >5 | Sensitivity = 0.79 Specificity = 0.91 Sensitivity = 0.84 Specificity = 0.91 | - | - | No |
| Zhao et al (2022) ⁶⁸ | Images of 200 individuals from the 3DBodyTex ⁶⁹ database | All | Accuracy = 0.71 | - | Accuracy = 0.88 | No |

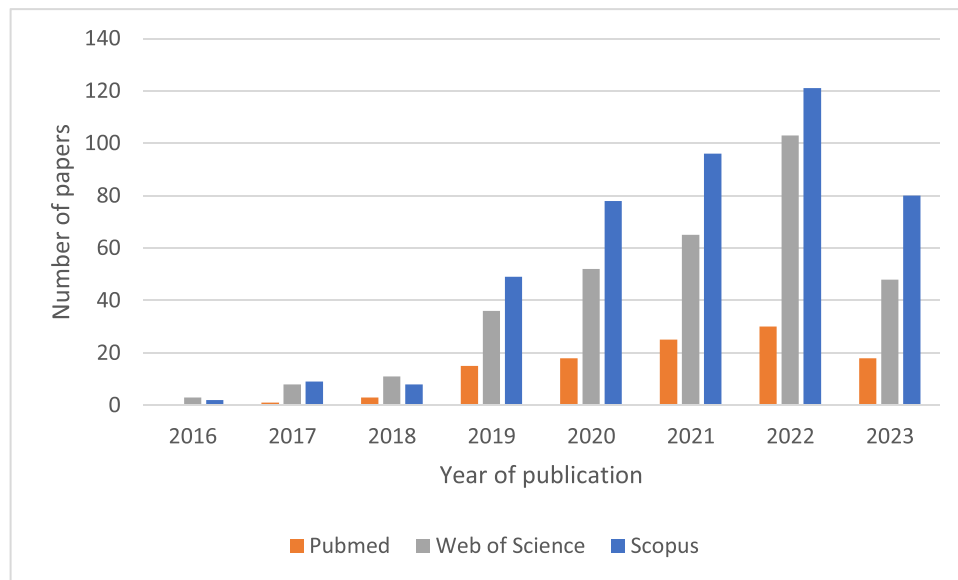


Fig. 3 The number of contributions indexed by PubMed,⁷¹ Scopus,⁷² and Web of Science⁷³ using the query “skin lesion classification CNN.” Data as of July 1, 2023 (not normalized).

allows for the early detection of skin changes. AI algorithms have been implemented in this system, but the manufacturer has not specified what kind of image analysis they perform, nor has the system performance been specified.

Some other TBS are described in a review paper.⁷⁰ The vast majority of them do not implement AI-based approaches. This study summarizes the current evidence on TBS for the early detection of melanoma. The systems analysis was conducted for clinical applications, including melanoma incidence, biopsy rates, and Breslow’s detected tumor index.

Discussion

Using ontologies and expert systems offers several advantages over traditional machine learning approaches. Whereas machine learning methods rely heavily on large data sets of verified images and training algorithms, ontologies provide a knowledge-driven approach incorporating expert insights and domain-specific information. Ontology-based expert systems support semantic annotation and inference capabilities to aid in the classification of lesion severity and the identification of potential skin cancers with high accuracy, with correct predictions exceeding 90%. By integrating expert knowledge and leveraging the semantic relationships defined in ontologies, expert systems can assist in decision-making processes, providing valuable insights for medical practitioners; however, due to the intensive development of AI methods using machine learning, in particular deep neural networks, expert systems constitute only a significant minority of the various skin analysis methods described in the literature.

Numerous examples of the classification of images depicting pathologic skin changes are found in the literature. A query for “skin lesion classification CNN” in the PubMed⁷¹ database returned 105 indexed contributions. The same query in Web of Science⁷² yielded 326 contributions, and in Scopus,⁷² there are 433. The increasing number of publications in all these databases indicates that interest in this research topic is growing yearly (Figure 3).

Unfortunately, some contributions relate to research conducted on relatively small local image data sets.⁷⁴⁻⁷⁶ Others involve experiments conducted on large data sets, but the details of the methods used are either not disclosed or are presented in a way making it impossible to repeat the experiment.⁷⁷⁻⁸¹ In other reports, serious methodologic errors have been made, such as excessive augmentation of selected classes followed by the division of images into training,³⁵ validation, and test sets. As a result, there may be similar images in each set, leading to overestimating the quantitative measures of classification accuracy.³⁵

Interesting publications

The most exciting publications are those in which the research was carried out on large data sets, and the image analysis method is presented in a way that allows the experiment to be repeated. The most objective skin lesion diagnosis results are available on ISIC challenge leaderboards, the most recent of which was published in 2020.¹⁰ Although the methods used have rarely been fully published in scientific journals. Summaries or short reports describing the methods used are available. Additionally, the full details of the procedures used in the training and test phases are often published together. The most effective techniques for diagnosing skin le-

sions employ an ensemble of classifiers based on deep convolutional networks.

Smartphone applications is promising in the early-stage diagnosis of skin pathologies. The introduction of AI into smartphone applications has increased the ability to search, track, and diagnose skin lesions. It is essential to detect and diagnose lesions early to achieve a high probability of stopping the development of pathologic changes.

The main advantages of the smartphone system approach include that the application can be widely distributed among medical personnel and patients and that the image database can be easily expanded, which is essential because deep learning artificial neural networks require a large amount of data to train efficiently. Such methods can be easily tested using cameras in smartphones.

There are also many more critical reviews concerning such solutions.^{55,82-86} These contributions draw attention to the numerous limitations of the applied methods, including the poor quality of the images taken by smartphones, errors due to photographs being taken in a different light, lack of standardization in taking photographs when less experienced people do this using their smartphones, bias in skin cancer detection by AI due to the data set it was trained on, and decreased visibility of lesions on darker skin.

The use of mobile dermatoscopes may increase their diagnostic value. One review⁵⁵ does not recommend the use of smartphone applications for skin lesion diagnosis due to the poor and variable performance of algorithms based on smartphones. The authors stated that these apps have not yet shown sufficient promise to warrant recommendation of their use. The current CE marking assessment processes are inadequate to protect patients against the risks created by smartphone diagnostic or risk-stratification apps.

All the aforementioned apps have a disclaimer that they cannot replace a dermatologic examination. They cannot provide the diagnosis and are not a substitute for regular dermatologist visits to date.

Total body systems may quickly and accurately analyze the patient's whole body to detect suspicious lesions. Due to a rapid examination, the system may be implemented for screening or epidemiologic studies, but TBS also have some limitations⁸⁷: 2% to 8% of melanomas are hypopigmented, and detection on a total body imaging system may be less reliable than with pigmented lesions. Also, AI systems are much less helpful in diagnosing rarer but necessary cutaneous and mucous membrane cancers such as trabecular carcinomas (Merkel cell cancers).

Total body systems face significant challenges related to the need to detect lesions in images covering a large part of the patient's body and matching pairs of lesions in followup examinations to detect new lesions or those that have changed or grown since the last examination. The relatively small number of systems^{60,61} also compromises the detection of neoplastic lesions. Automatic analysis of images from TBS systems is more complex than evaluating images of individual moles obtained from a dermatoscope or a digital

camera. This is due to the small size of the imaged lesions and geometric distortions caused by the complex shapes of the human body, which results in not all moles being perpendicular to the camera axis. The AI algorithms used in such systems provide a mole detection sensitivity of 0.9 and a matching accuracy of 0.8. These values are promising; however, the methods described in the literature have been validated on a relatively small number of patients (several dozen people, several tens of thousands of analyzed nevi). These systems (except the system described in the study⁵⁹) have also not been clinically validated.

Diagnosis of skin lesions

This review of various types of systems supporting the diagnosis of skin lesions shows that the AI methods are practical tools. Of particular note are the different deep network architectures that ensure the exact operation of such systems. This is confirmed by the obtained values of measures evaluating the effectiveness of the applied algorithms. Many of these algorithms have been validated on large data sets, increasing the credibility of their results. The implementation of AI may also reduce the costs of melanoma screening. As was demonstrated in 1988, the cost of diabetic retinopathy screening can be lower if AI algorithms are used instead of expert judgment.⁸⁸ A similar situation may occur in dermatology, in which screening tests require many imaging analyses performed by dermatologists or other health providers. Using AI in dermatology, particularly in image-based dermatologic diagnosis, may bring modern solutions and benefits to patients and improve diagnosing by making it more objective and reliable.

Artificial intelligence methods pose limitations and threats that apply to dermatology and other fields of medicine. One such limitation of using AI in therapy stems from challenges such as small and inadequately described training data sets, leading to limited accuracy and practicality in clinical applications. Relevant data and large, well-described image data sets are necessary for practical AI algorithm training; otherwise, AI algorithms may not obtain sufficient generalization, which will translate into wrong decisions (eg, when analyzing images from different medical centers). There is a risk of misdiagnosis or inaccurate results, especially when attempting to diagnose diseases with minimally available training examples. Image artifacts and a lack of unified disease identification criteria hinder the recognition and diagnosis of AI in dermatology. Addressing these limitations—including the shortage of rare disease cases of patients with rare skin cancer-prone genodermatoses, such as albinism and xeroderma pigmentosum, and patients who are cancer-prone due to treatment, such as solid organ transplant recipients—remains a significant challenge for implementing AI in dermatology.⁸⁹ Another limitation may be the incorrect methodology for using AI algorithms (eg, excessive data augmentation, incorrect validation of algorithms). There

is also a risk that undue reliance on AI may lead to the decline of traditional diagnostic and treatment methods among dermatologists, especially the younger and less experienced. In the field of dermatology, future AI trends include increasing support for diagnostic platforms, the emergence of intelligent medical devices and instruments, the utilization of transmedical intelligent medical equipment, integration with 5G networks, the implementation of cloud-based intelligent health care models, and the popularization of AI consultations⁹⁰; however, AI is not expected to replace dermatologists, as it lacks certain human qualities completely, and physicians will continue to play a vital role in patient care. Proper understanding and utilization of AI can provide convenience to physicians and improve the quality of services for patients in dermatology.

Legal issues

The use of AI in medicine also faces possible legal issues. The EU has legislative and policy initiatives relating to using such algorithms and recognizing them as typical clinical procedures.⁹¹ Currently, a new AI liability legal framework is in development.⁹² The United States seems to be more advanced in this regard, where the Food and Drug Administration (FDA), in collaboration with the International Medical Device Regulators Forum, has developed a certification path for medical devices implementing AI algorithms and the algorithms themselves.⁹³ Accepted solutions in the software category as a medical device can be used in clinical practice. Currently, the FDA has approved more than 500 such algorithms,⁹⁴ but none are suitable for dermatology (the same applies to the EU-supported algorithms). Currently, solutions intended for applications mainly in radiology and cardiology dominate, and to a lesser extent in neurology, gastroenterology and hematology. Appropriate legal regulations are also needed to enable patient medical data exchange between clinical centers and other scientific institutions to develop high-quality AI algorithms. Another aspect that should be considered when using AI is ethical issues, such as possibly informing and obtaining informed consent from patients for this diagnostic procedure. With further development of AI methods, as the level of their “awareness” and “decisiveness” increases, a vital issue to be resolved will be the issue of legal liability for any diagnostic errors made by the algorithms.

Conclusions

The contribution and role of AI in supporting dermatologic diagnostics are growing; a similar conclusion applies to other medical specialties. Developing these algorithms allows us to obtain results comparable to the assessment of specialists. In this context, AI emerges as a valuable tool that aids

physicians, provided the algorithms are adequately learned and reliably validated.

The effective integration of AI in diagnostics demands physicians to possess a suitable level of awareness and knowledge, often at the intersection of medicine, biomedical engineering, and informatics. They need to understand the limitations of the algorithms in use and the rationale behind diagnostic decisions facilitated by explainable AI. Consequently, clinical engineers are increasingly responsible for maintaining and developing these algorithms, including scenarios in which updates and retraining of neural networks are necessary.

Many issues about using AI in clinical practice remain to be resolved. These issues are technical (ie, ensuring access to a large amount of medical data of appropriate quality, which is necessary for the practical training of most of the algorithms used) and legal (ie, developing consistent rules and conditions enabling AI algorithms in the clinical setting).

Declaration of competing interest

The authors declare no conflicts of interest.

References

1. Marks R. Epidemiology of melanoma. *Clin Exp Dermatol*. 2000;25:459–463.
2. Matthews NH, Li W-Q, Qureshi AA, Weinstock MA, Cho E. Epidemiology of melanoma. *Exon Publications*. 2017;6:3–22.
3. Tsao H, Olazagasti JM, Cordero KM, et al. Early detection of melanoma: reviewing the ABCDEs. *J Am Acad Dermatol*. 2015;72:717–723.
4. Gowthami V, Sneha G. Melanoma detection using recurrent neural network. *Lecture Notes Electr Eng*. 2021;700:1563–1573.
5. Zhang B, Wang Z, Gao J, et al. Short-term lesion change detection for melanoma screening with novel Siamese neural network. *IEEE Trans Med Imaging*. 2021;40:840–851.
6. Abu-Nasser B. Medical expert systems survey. *Int J Soft Comp Math Control*. 2017;4:25–37.
7. Tschandl P, Rinner C, Apalla Z, et al. Human–computer collaboration for skin cancer recognition. *Nat Med*. 2020;26:1229–1234.
8. Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci Data*. 2018;5:1–9.
9. International Skin Imaging Collaboration. Available at: <https://www.isic-archive.com/>. Accessed June 24, 2023.
10. Zawacki A, Helba B, Shih G, et al. SIIM-ISIC melanoma classification. Available at: <https://www.kaggle.com/c/siim-isic-melanoma-classification/overview>. Accessed July 1, 2023.
11. Gutman D, F Codella NC, Celebi E, et al. Skin lesion analysis toward melanoma detection: a challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC). *arXiv*. 2016:1605.01397.
12. Codella NCF, Gutman D, Celebi ME, et al. Skin lesion analysis toward melanoma detection: a challenge at the 2017 International symposium on biomedical imaging (ISBI), hosted by the International Skin Imaging Collaboration (ISIC). *Proc Int Symp Biomed Imaging*. 2018:168–172.
13. Codella N, Rotemberg V, Tschandl P, et al. Skin lesion analysis toward melanoma detection 2018: a challenge hosted by the International Skin Imaging Collaboration (ISIC). *arXiv*. 2019:1902.03368.

14. Combalia M, Codella NCF, Rotemberg V, et al. BCN20000: dermoscopic lesions in the wild. *arXiv*. 2019:1908.02288.
15. Rotemberg V, Kurtansky N, Betz-Stablein B, et al. A patient-centric dataset of images and metadata for identifying melanomas using clinical context. *Sci Data*. 2021;8:34.
16. International Skin Imaging Collaboration (ISIC). The ISIC 2020 Challenge Dataset. Available at: <https://challenge2020.isic-archive.com/>. Accessed June 28, 2023.
17. DermNet. Dermatology resource. Available at: <https://dermnetnz.org/>. Accessed July 8, 2023.
18. University of Edinburgh. Dermofit Image Library. Available at: <https://licensing.edinburgh-innovations.ed.ac.uk/product/dermofit-image-library>. Accessed July 8, 2023.
19. Maragoudakis M, Maglogiannis I. A medical ontology for intelligent web-based skin lesions image retrieval. *Health Informatics J*. 2011;17:140–157.
20. Abbes W, Sellami D. Automatic skin lesions classification using ontology-based semantic analysis of optical standard images. *Procedia Comput Sci*. 2017;112:2096–2105.
21. Laskaris N, Ballerini N, Fisher RB, Aldridge B, Rees J. Fuzzy description of skin lesions. *SPIE Med Imaging*. 2010;7627:377–386.
22. Abbes W, Sellami D, Marc-Zwecker S, Zanni-Merk C. Fuzzy decision ontology for melanoma diagnosis using KNN classifier. *Multimed Tools Appl*. 2021;80:25517–25538.
23. Tang P, Liang Q, Yan X, Xiang S, Zhang D. GP-CNN-DTEL: Global-Part CNN model with data-transformed ensemble learning for skin lesion classification. *IEEE J Biomed Health Inform*. 2020;24:2870–2882.
24. Olayah F, Senan EM, Ahmed IA, Awaji B. AI techniques of dermoscopy image analysis for the early detection of skin lesions based on combined CNN features. *Diagnostics*. 2023;13:1314.
25. Ahmed SAA, Yanikoglu B, Goksu O, Aptoula E. Skin lesion classification with deep CNN ensembles. *2020 28th Signal Processing and Communications Applications Conference (SIU)*. Gaziantep, Turkey: IEEE; 2020:1–4.
26. Russakovsky O, Deng J, Su H, et al. ImageNet large scale visual recognition challenge. *Int J Comput Vis*. 2015;115:211–252.
27. Saba T, Khan MA, Rehman A, Marie-Sainte SL. Region extraction and classification of skin cancer: a heterogeneous framework of deep CNN features fusion and reduction. *J Med Syst*. 2019;43:1–19.
28. Villa-Pulgarin JP, Ruales-Torres AA, Arias-Garzón D, et al. Optimized convolutional neural network models for skin lesion classification. *Comp Mater Continua*. 2021;70:2131–2148.
29. Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions. *Computer Vision and Pattern Recognition (CVPR)*. Boston, MA, USA: IEEE; 2015:1–9.
30. Kassem MA, Hosny KM, Fouad MM. Skin lesions classification into eight classes for ISIC 2019 using deep convolutional neural network and transfer learning. *IEEE Access*. 2020;8:114822–114832.
31. Hasan MK, Elahi MTE, Alam MA, Jawad MT, Martí R. DermoExpert: skin lesion classification using a hybrid convolutional neural network through segmentation, transfer learning, and augmentation. *Inform Med Unlocked*. 2022;28:100819.
32. Hasan MK, Dahal L, Samarakoon PN, Tushar FI, Martí R. DSNet: automatic dermoscopic skin lesion segmentation. *Comput Biol Med*. 2020;120:103738.
33. Ha Q, Liu B, Liu F. Identifying melanoma images using EfficientNet ensemble: winning solution to the SIIM-ISIC melanoma classification challenge. *arXiv*. 2020:1908.02288.
34. Ha Q, Liu B, Liu F. haqishen/SIIM-ISIC Melanoma Classification 1st Place Solution. Available at: <https://github.com/haqishen/SIIM-ISIC-Melanoma-Classification-1st-Place-Solution>. Accessed July 1, 2023.
35. Aldhyani THH, Verma A, Al-Adhaileh MH, et al. Multi-class skin lesion classification using a lightweight dynamic kernel deep-learning-based convolutional neural network. *Diagnostics*. 2022;12:2048.
36. Winkler JK, Blum A, Kommos K, et al. Assessment of diagnostic performance of dermatologists cooperating with a convolutional neural network in a prospective clinical study: human with machine. *JAMA Dermatol*. 2023;159:621–627.
37. FotoFinder. AI Moleanalyzer Pro. Available at: <https://www.fotofinder.de/en/technology/artificial-intelligence/moleanalyzer-pro>. Accessed July 8, 2023.
38. DermLite. DermLite adapters for iPhone, any Android smartphone, iPad, or tablet. Available at: <https://dermlite.com/collections/connection-kits>. Accessed June 29, 2023.
39. IDCPC. Molescope. Available at: <https://www.idcp.eu/molescope-eu>. Accessed June 29, 2023.
40. Alves J, Moreira D, Alves P, Rosado L, João M, Vasconcelos M. Automatic focus assessment on dermoscopic images acquired with smartphones. *Sensors*. 2019;19:4957.
41. Yilmaz A, Gencoglan G, Varol R, Demircali AA, Keshavarz M, Uvet H. MobileSkin: classification of skin lesion images acquired using mobile phone-attached hand-held dermoscopes. *J Clin Med*. 2022;11:5102.
42. Jaworek-Korjakowska J, Kleczek P. ESkin: study on the smartphone application for early detection of malignant melanoma. *Wirel Commun Mob Comput*. 2018;2018:5767630.
43. Kalwa U, Legner C, Kong T, Pandey S. Skin cancer diagnostics with an all-inclusive smartphone application. *Symmetry*. 2019;11:790.
44. Mendonca T, Ferreira PM, Marques JS, Marcal ARS, Rozeira J. PH2: a dermoscopic image database for research and benchmarking. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Osaka, Japan: IEEE; 2013:5437–5440.
45. Udrea A, Mitra GD, Costea D, et al. Accuracy of a smartphone application for triage of skin lesions based on machine learning algorithms. *Wiley Online Library*. 2019;34:648–655.
46. SkinVision. Skin cancer melanoma detection app. Available at: <https://www.skinvision.com/>. Accessed July 8, 2023.
47. Francese R, Frasca M, Risi M, Tortora G. A mobile augmented reality application for supporting real-time skin lesion analysis based on deep learning. *J Real Time Image Process*. 2021;18:1247–1259.
48. Krohling B, Castro PBC, Pacheco AGC, Krohling RA. A smartphone based application for skin cancer classification using deep learning with clinical images and lesion information. *arXiv*. 2021:2104.14353.
49. Pacheco AGC, Krohling RA. The impact of patient clinical information on automated skin cancer detection. *Comput Biol Med*. 2020;116:103545.
50. Maduranga P, Nandasena D, Maduranga M. Mobile-based skin disease diagnosis system using convolutional neural networks (CNN). *IJ Image Graphics Signal Process*. 2022;3:47–57.
51. Kränke T, Tripolt-Droschl K, Röd L, Hofmann-Wellenhof R, Kopitz M, Tripolt M. New AI-algorithms on smartphones to detect skin cancer in a clinical setting: a validation study. *PLoS One*. 2023;18:e0280670.
52. SkinScreener. An app to assess the risk of skin lesions. Available at: <https://skinscreener.com/en>. Accessed July 8, 2023.
53. Thissen M, Udrea A, Hacking M, Von Braunmühl T, Ruzicka T. MHealth app for risk assessment of pigmented and nonpigmented skin lesions: a study on sensitivity and specificity in detecting malignancy. *Telemed e-Health*. 2017;23:948–954.
54. de Carvalho TM, Noels E, Wakkee M, Udrea A, Nijsten T. Development of smartphone apps for skin cancer risk assessment: progress and promise. *JMIR Dermatol*. 2019;2:e13376.
55. Freeman K, Dinnes J, Chuchu N, et al. Algorithm based smartphone apps to assess risk of skin cancer in adults: systematic review of diagnostic accuracy studies. *BMJ*. 2020;368:m127.
56. Taeg SC, Freeman WT, Tsao H. A reliable skin mole localization scheme. *Proc IEEE Int Conf Comp Vision*; 2007:1–8.
57. Bogo F, Romero J, Peserico E, Black MJ. Automated detection of new or evolving melanocytic lesions using a 3D body model. *Lecture Notes Comp Sci*. 2014;8673:593–600.
58. Korotkov K, Quintana J, Puig S, Malveyh J, Garcia R. A new total body scanning system for automatic change detection in multiple pigmented skin lesions. *IEEE Trans Med Imaging*. 2015;34:317–338.

59. Korotkov K, Quintana J, Campos R, et al. An improved skin lesion matching scheme in total body photography. *IEEE J Biomed Health Inform.* 2019;23:586–598.
60. Birkenfeld JS, Tucker-Schwartz JM, Soenksen LR, Avilés-Izquierdo JA, Marti-Fuster B. Computer-aided classification of suspicious pigmented lesions using wide-field images. *Comput Methods Programs Biomed.* 2020;195:105631.
61. Soenksen LR, Kassis T, Conover ST, et al. Using deep learning for dermatologist-level detection of suspicious pigmented skin lesions from wide-field images. *Sci Transl Med.* 2021;13:17.
62. Strzelecki MH, Strąkowska M, Kozłowski M, Urbańczyk T, Wielowieyska-Szybińska D, Kociołek M. Skin lesion detection algorithms in whole body images. *Sensors.* 2021;21:6639.
63. Szczypiński PM, Sprawka K. Orthorectification of skin nevi images by means of 3D model of the human body. *Sensors.* 2021;21:8367.
64. Strąkowska M, Kociołek M. Skin lesion matching algorithm for application in full body imaging systems. In: Pietka E, Badura P, Kawa J, Wicławek W, eds. *Information Technology in Biomedicine. ITIB 2022. Advances in Intelligent Systems and Computing, Vol. 1429.* Cham: Springer; 2022:222–233.
65. Skopia Estetic Clinic. Dermo Studio. Available at: <https://dermo-studio.eu/#kontakt>. Accessed July 1, 2023.
66. Betz-Stablein B, D'Alessandro B, Koh U, et al. Reproducible naevus counts using 3D total body photography and convolutional neural networks. *Dermatology.* 2022;238:4–11.
67. Koh U, Janda M, Aitken JF, et al. Mind Your Moles" study: protocol of a prospective cohort study of melanocytic naevi. *BMJ Open.* 2018;8:e025857.
68. Zhao M, Kawahara J, Abhishek K, Shamanian S, Hamarneh G. Skin3D: detection and longitudinal tracking of pigmented skin lesions in 3D total-body textured meshes. *Med Image Anal.* 2022;77:102329.
69. Saint A, Ahmed E, Shabayek AER, et al. 3DBodyTex: textured 3D body dataset. *Proc Int Conf 3D Vision;* 2018:495–504.
70. Hornung A, Steeb T, Wessely A, et al. The value of total body photography for the early detection of melanoma: a systematic review. *Int J Environ Res Public Health.* 2021;18:1726.
71. PubMed. National Library of Medicine. Available at: <https://pubmed.ncbi.nlm.nih.gov/>. Accessed July 8, 2023.
72. Scopus. Elsevier B.V. Available at: <https://www.scopus.com/search/form.uri?display=basic#basic>. Accessed July 8, 2023.
73. Web of Science Core Collection. Clarivate. Available at: <https://www.webofscience.com/wos/woscc/basic-search>. Accessed July 8, 2023.
74. Pramanik R, Banerjee B, Efimenko G, Kaplun D, Sarkar R. Monkeypox detection from skin lesion images using an amalgamation of CNN models aided with beta function-based normalization scheme. *PLoS One.* 2023;18:e0281815.
75. Goceri E, Karakas AA. Comparative evaluations of CNN based networks for skin lesion classification. *The 14th International Conference on Computer Graphics, Visualization, Computer Vision, And Image Processing (CGVCVIP 2020), Zagreb, Croatia;* 2020:1–6.
76. Lihacova I, Bondarenko A, Chizhov Y, et al. Multi-class CNN for classification of multispectral and autofluorescence skin lesion clinical images. *J Clin Med.* 2022;11:2833.
77. Nersisson R, Iyer TJ, Joseph Raj AN, Rajangam V. A dermoscopic skin lesion classification technique using YOLO-CNN and traditional feature model. *Arab J Sci Eng.* 2021;46:9797–9808.
78. Alfaro E, Fonseca XB, Albornoz EM, Martinez CE, Ramirez SC. A brief analysis of U-Net and Mask R-CNN for skin lesion segmentation. *2019 IEEE International Work Conference on Bioinspired Intelligence (IWOBI).* Budapest, Hungary: IEEE; 2019:123–126.
79. Dhar P, Guha S. Engineering and manufacturing, skin lesion detection using fuzzy approach and classification with CNN. *I J Eng Manufact.* 2021;1:11–18.
80. Kumar RS, Singh A, Srinath VS, Arasu V, Thomas NK. Machine learning on web: skin lesion classification using CNN. *2022 International Conference on Inventive Computation Technologies (ICICT).* Lalitpur, Nepal: IEEE; 2022:260–265.
81. Sankar Raja Sekhar K, Ranga Babu T, Prathibha G, Vijay K, Chiau Ming L. Dermoscopic image classification using CNN with handcrafted features. *J King Saud Univ Sci.* 2021;33:101550.
82. Beltrami EJ, Brown AC, Salmon PJM, Leffell DJ, Ko JM, Grant-Kels JM. Artificial intelligence in the detection of skin cancer. *J Am Acad Dermatol.* 2022;87:1336–1342.
83. Matin RN, Dinnes J. AI-based smartphone apps for risk assessment of skin cancer need more evaluation and better regulation. *Br J Cancer.* 2021;124:1749–1750.
84. Ouellette S, Rao BK. Usefulness of smartphones in dermatology: a US-based review. *J Environ Res Public Health.* 2022;19:3553.
85. Rat C, Hild S, Sérandour JR, et al. Use of smartphones for early detection of melanoma: systematic review. *J Med Internet Res.* 2018;20:e135.
86. Xiong M, Pfau J, Young AT, Wei ML. Artificial intelligence in teledermatology. *Curr Dermatol Rep.* 2019;8:85–90.
87. Rayner JE, Laino AM, Nufer KL, et al. Clinical perspective of 3D total body photography for early detection and screening of melanoma. *Front Med (Lausanne).* 2018;5:371716.
88. Grzybowski A, Singhanetr P, Nanegrungsunk O, Ruamviboonsuk P. Artificial intelligence for diabetic retinopathy screening using color retinal photographs: from development to deployment. *Ophthalmol Ther.* 2023;12:1419–1437.
89. Steele L, Velazquez-Pimentel D, Thomas BR. Do AI models recognise rare, aggressive skin cancers? An assessment of a direct-to-consumer application in the diagnosis of Merkel cell carcinoma and amelanotic melanoma. *J Eur Acad Dermatol Venereol.* 2021;35:e877–e879.
90. Li Z, Koban KC, Schenck TL, Giunta RE, Li Q, Sun Y. Artificial intelligence in dermatology image analysis: current developments and future trends. *J Clin Med.* 2022;11:6826.
91. Artificial Intelligence and Medical Devices Regulation. Discussing the legal framework and the ethical challenges within the CORE-MD project. Available at: <https://www.core-md.eu/artificial-intelligence-and-medical-devices-regulation-discussing-the-legal-framework-and-the-ethical-challenges-within-the-core-md-project/>. Accessed July 8, 2023.
92. European Commission. Liability rules for artificial intelligence. Available at: https://commission.europa.eu/business-economy-euro/doing-business-eu/contract-rules/digital-contracts/liability-rules-artificial-intelligence_en. Accessed July 8, 2023.
93. Food and Drug Administration. Global approach to software as a medical device. Available at: <https://www.fda.gov/medical-devices/software-medical-device-samd/global-approach-software-medical-device>. Accessed July 3, 2023.
94. Food and Drug Administration. Artificial intelligence and machine learning (AI/ML)-enabled medical devices. Available at: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices>. Accessed July 3, 2023.