Received: 24 January 2022 | Accepted: 5 January 2023

DOI: 10.1111/jdv.18963

JEADV JOURNAL OF THE EUROPEAN ACADEMY OF DERMATOLOGY & VENEREOLOGY

SYSTEMATIC REVIEW

The application of artificial intelligence in the detection of basal cell carcinoma: A systematic review

Y. Widaatalla[1] | T. Wolswijk[2] | F. Adan[2] | L. M. Hillen[3] | H. C. Woodruff[1,4] | I. Halilaj[1] | A. Ibrahim[5] | P. Lambin[1,4] | K. Mosterd[2]

[1]The D-Lab, Department of Precision Medicine, GROW School for Oncology and Reproduction, Maastricht University, Maastricht, The Netherlands

[2]Department of Dermatology, GROW School for Oncology and Reproduction, Maastricht University Medical Center+, Maastricht, The Netherlands

[3]Department of Pathology, GROW School for Oncology and Reproduction, Maastricht University Medical Center+, Maastricht, The Netherlands

[4]Department of Radiology and Nuclear Medicine, Maastricht University Medical Center+, Maastricht, The Netherlands

[5]Department of Radiology, Columbia University Irving Medical Center, New York, New York, USA

Correspondence
Y. Widaatalla, Department of Precision Medicine, Maastricht University, Duboisdomein 30, Maastricht, 6229 GT, The Netherlands.
Email: y.widaatalla@maastrichtuniversity.nl

Abstract

Basal cell carcinoma (BCC) is one of the most common types of cancer. The growing incidence worldwide and the need for fast, reliable and less invasive diagnostic techniques make a strong case for the application of different artificial intelligence techniques for detecting and classifying BCC and its subtypes. We report on the current evidence regarding the application of handcrafted and deep radiomics models used for the detection and classification of BCC in dermoscopy, optical coherence tomography and reflectance confocal microscopy. We reviewed all the articles that were published in the last 10 years in PubMed, Web of Science and EMBASE, and we found 15 articles that met the inclusion criteria. We included articles that are original, written in English, focussing on automated BCC detection in our target modalities and published within the last 10 years in the field of dermatology. The outcomes from the selected publications are presented in three categories depending on the imaging modality and to allow for comparison. The majority of articles (n = 12) presented different AI solutions for the detection and/or classification of BCC in dermoscopy images. The rest of the publications presented AI solutions in OCT images (n = 2) and RCM (n = 1). In addition, we provide future directions for the application of these techniques for the detection of BCC. In conclusion, the reviewed publications demonstrate the potential benefit of AI in the detection of BCC in dermoscopy, OCT and RCM.

INTRODUCTION

Basal cell carcinoma (BCC) is one of the most common forms of cancer,[1-3] with a rising incidence worldwide. To date, histopathological examination of a punch biopsy is the gold standard to distinguish BCC from alternative diagnoses and to determine the BCC subtype.[1,4] However, a punch biopsy is an invasive procedure, with risks of pain and bleeding during the procedure and the additional chance of infection and/or scarring.[5] Moreover, awaiting histopathological assessment may be stressful for many patients. Considering the rising incidence of BCC, which causes a major burden on healthcare systems, finding alternatives to an invasive biopsy is desired.

The last century has seen many advances in medical imaging from the first X-ray to more sophisticated imaging techniques, such as computed tomography (CT) and magnetic resonance imaging (MRI).[6] This progress, concurrent with advances in computational power and machine learning techniques, often termed artificial intelligence (AI) in the literature, eased the transformation of medical images into quantitative minable data that could be used to build diagnostic, predictive and prognostic clinical decision support systems (cDSS).[7,8]

Quantitative imaging analysis techniques that are being extensively investigated to develop cDSS include handcrafted radiomics and deep learning. Handcrafted radiomic features

P. Lambin and K. Mosterd: Senior authors contributed equally.

(HRFs) are mathematical formulas applied to the array of intensity values representing the medical image. There are different groups of HRFs, which give insight into the distribution of pixel intensities (first-order features), the relationship between pixels intensities (textures), complex relations between pixel values and position (higher order features) and many more.[9] These HRFs can be further correlated univariably or multivariably with clinical outcomes or biological characteristics using machine learning techniques.

Deep learning (DL) is a branch of machine learning that is data-driven and uses neural network-like structures to mimic the structure of the human brain. A neural network is a layered-like structure, consisting of an input, hidden layer(s) and an output layer.[10] These networks can be trained on past data to make predictions about unseen, new data and are considered 'deep' when it contains more than one hidden layer. Unlike HRFs, quantitative features extracted for DL purposes are less well-understood, leading to the term 'black box' to be applied to such models.[11] Trained networks are used to recognize patterns within the image to perform different tasks, such as segmentation, classification and prediction.[12] Hence, it is no surprise that DL is gaining increased popularity for medical imaging analysis including dermatological images for various skin conditions, such as BCC which is the focus of this study.

In this systematic review, we focus on three non-invasive imaging modalities used for diagnosing BCC: (i) dermoscopy, a routine step in diagnosing BCC; (ii) optical coherence tomography (OCT), a relatively new tool using light interferometry to provide morphological insight similar to histopathology; and (iii) reflectance confocal microscopy (RCM), which provides high-resolution transversal images. We report on the published state-of-the-art AI methods used for automated detection, localization and classification of BCC, their challenges and limitations, as well as future perspectives for the clinical use of these techniques in the diagnosis of BCC.

DERMATOLOGICAL ASSESSMENT AND IMAGING TOOLS

Dermoscopic images

Dermoscopy (epiluminescence microscopy) is a non-invasive diagnostic technique, developed to assist the dermatologist in the identification of pigmented skin lesions.[13] It helps with the visualization of the lesion beyond the abilities of the naked eye using a hand-held microscope device, sometimes in conjunction with an immersion fluid to decrease reflections and to make the skin layer more permeable to light (Figure 1a).[14]

A few studies investigated the diagnostic accuracy before and after the introduction of dermoscopy. For instance, Braun et al.[15] reported a 20% increase in the diagnostic accuracy of pigmented lesions (from 75% to 95%). Furthermore, Reiter et al. reviewed 17 studies comparing the sensitivity and specificity in diagnosing BCC with and without dermoscopy and reported that the overall sensitivity and specificity improved by 18% and 1%, respectively.[16]

The diagnostic accuracy can still be suboptimal when dermoscopy is performed and evaluated by an inexperienced dermatologist.[17] Moreover, dermoscopic images are subjected to intra- and interobserver variabilities. Ultimately, this visualization/interpretation subjectivity issue together with the advancements of machine learning methods paved the road for exploring computerized techniques. This also potentially reduces the lack of reproducibility induced by the subjectivity of human interpretation.[18]

OCT images

By the late 1990s, OCT imaging was introduced to dermatology practices.[19] The core concept of OCT is similar to ultrasound imaging as it measures the echo delay and intensity produced by a reflected signal, in this case infrared light instead of sound.[20] OCT acquires real-time cross-sectional images of the skin with a penetration depth of approximately 1.0–1.5 mm and a <7.5 μm lateral and <5 μm axial optical resolution (Figure 1b), although new techniques may increase penetration depth.[21]

Based on the reflections, architectural details of lesions and tissues can be visualized enabling the identification of a lesion as BCC and its respective subtype (superficial, nodular or infiltrative). Hussain et al.[22] established a set of morphological features such as a honeycomb-like structure, disruption of the dermal–epidermal junction, hypo reflective lateral border and dilated vasculature. Sinx et al.[23] investigated the added diagnostic value of OCT in diagnosing BCC and its subtypes. The prospective cohort included 250

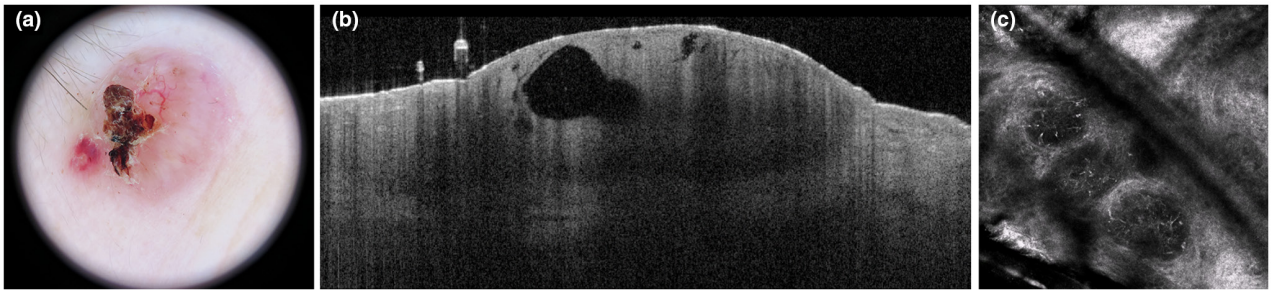

FIGURE 1 Showing a BCC nest: (a) Dermoscopy image of a BCC. (b) OCT scan showing a BCC nest. (c) RCM image showing a BCC nest.

suspicious lesions requiring a biopsy to confirm the diagnosis of BCC. The accuracy of clinical examination alone was compared with the diagnosis based on OCT in conjunction with a clinical photograph of the lesion. The results showed a higher area under the curve (AUC) of 91.2% for the OCT combined with the clinical photograph compared to 85.6% for the clinical and dermoscopic examination alone in discriminating BCC from non-BCC lesions. In addition, Adan et al.[24] investigated the percentage of biopsies that could be omitted by relying only on OCT. The results showed that 66% of punch biopsies could be omitted, since the diagnosis of BCC and its subtype could be made with high confidence on OCT. Moreover, compared with regular care (i.e. biopsy), OCT was cost-effective and noninferior regarding the effectiveness of the total diagnostic and therapeutic strategy. If OCT analysis leads to a high-confidence diagnosis of BCC and its subtype, a patient could potentially receive a diagnosis and a treatment within a single consultation. Efficiency is thus increased by saving time, costs and finally yet importantly, it may improve the patient's well-being during the diagnostic procedure.

Reflectance confocal microscopy

Reflectance confocal microscopy is a non-invasive imaging technique that allows for the visualization of cellular and subcellular structures of skin.[25] This imaging modality works by selectively collecting the laser light reflected from the specimen horizontally, allowing the light source to scan the area (skin) at a fixed depth with high resolution using a two-dimensional grid. Furthermore, confocal microscopy captures multiple, in-real two-dimensional images (Figure 1c) that can be stacked allowing for the reconstruction of three-dimensional images.[26]

Nori et al.[27] investigated the sensitivity and specificity of RCM in diagnosing BCC. The data set consisted of a total of 152 benign and malignant lesions. The authors used histological criteria that correlate with the presence of BCC in RCM, such as the presence of elongated monomorphic nuclei, increased vasculature and prominent inflammatory infiltration. The authors reported a sensitivity and specificity of 83% and 95%, respectively, when four or more histological criteria are present.

MATERIALS AND METHODS

Search and databases

In July 2022, we searched PubMed, Web of Science and EMBASE using the combinations of the following keywords ('Basal cell carcinoma' OR 'Skin cancer' OR 'Skin neoplasms' OR 'Non-melanoma skin cancer') AND ('Deep learning' OR 'Computer-aided diagnosis' OR 'Machine learning' OR 'Automated detection') AND ('Optical Coherence Tomography' OR 'Dermoscopy' OR 'Dermatoscopy' OR 'reflectance confocal microscopy').

We included articles that are original, written in English, focussing on automated BCC detection in our target modalities and published within the last 10 years in the field of dermatology. We excluded reviews, editorials, opinion papers, non-English studies, studies with no distinction between BCC and other nonmelanoma skin cancers (NMSC), studies with invasive sampling, studies with nonoriginal methodology and studies that do not provide a clear metric for BCC detection.

In total, 883 publications were exported from the three databases to the EndNote citation manager.[28] Then, 123 duplicates were removed leaving 760 articles. Based on the title, we removed 666 articles. Furthermore, we removed 51 publications based on reading their abstracts. Furthermore, we excluded 28 articles based on the following reasons, nonoriginal methodology (n = 5); invasive diagnostic modality (n = 4); BCC accuracy was ancillary (n = 9); different imaging modality (n = 2); nonhuman studies (n = 1); no diagnostic parameter is provided (n = 1); inaccessible manuscript (n = 3); and conference papers (n = 3). Consensus was achieved between the two reviewers (Y.W and T.W) in each mentioned step in the inclusion and exclusion process. The publications selection was performed in accordance with the PRISMA statement,[29] and 15 articles were selected for the systematic review (Figure 2). Furthermore, the selected articles were divided into three categories according to the imaging modality.

This study reports on the classification performance of the selected studies using three main metrics, accuracy, sensitivity and specificity when these were available or could be derived. Otherwise, the AUC or detection rate is reported. Accuracy is defined by the total number of correct predictions made by the classifier, divided by the total number of predictions ((true positive + true negative)/(true positive + true negative + false positive + false negative)). It is ideally used in data sets with a balanced class distribution. In the case of multiclass classifiers, with unequal class distribution, we calculated the sensitivity and specificity for the specific class (BCC). Sensitivity refers to the ability of the classifier to correctly predict the positive class out of the total number of the positive class. While specificity refers to the ability of the classifier to predict correctly the negative class out of the total number of the negative class.

RESULTS

BCC detection in Dermoscopy images

Twelve articles reported on the detection of BCC in dermoscopy images.[30-41] Half of the studies included used retrospectively collected datasets, such as ISIC-2018 and ISIC-2019, which contain multiple skin conditions including malignant lesions such as BCC and benign conditions.[42,43] The

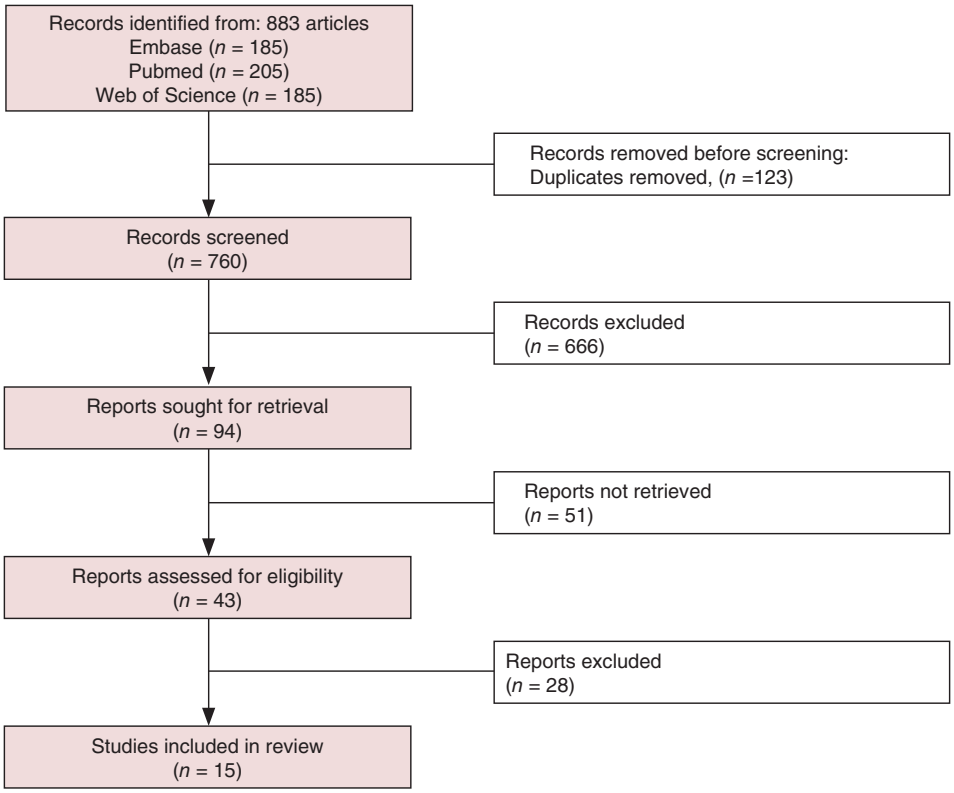

FIGURE 2 Articles selection process.

number of images used to train the model varied between studies, with only one publication using less than 1000 images to train and test the model.[34] Two articles used more than 22,000 images, whereas the rest used between 1000 and 20,000 images.[31,33]

The performance of the algorithms was reported using different outcomes, with sensitivity and specificity being the primary metric in eight studies, while AUC was reported in four studies (Table 1). Two articles reported on the accuracy for multiple classes; hence, we calculated sensitivity and specificity from the provided confusion matrix (Table 1).[30,31] The accuracy (for binary outcome) was reported in one study,[37] while one study[39] reported the detection rate as the main outcome. Two studies[32,34] developed a binary classier (BCC vs no-BCC) using relatively comparable data set sizes. However, both used different training methods; Kefel et al. used logistic regression for classification, while Kharazmi et al. utilized a spare autoencoder to achieve the same task. When comparing the performance, Kharazmi et al. and Kefel et al. reported different results with AUCs of 0.91 and 0.88, respectively (Table 1).

Two studies[35,37] approached BCC detection in dermoscopy images by segmenting the vascular structures (telangiectasia). Kharazmi et al. used an unsupervized learning algorithm (K-means clustering) to categorize unlabelled data depending on feature similarity. Finally, a random forest classifier was used to differentiate benign lesions from BCC. The proposed method achieved an AUC of 0.97. Maurya et al. used U-Net, which is a deep neural network often used for pixel-based segmentation tasks. The proposed method achieved an accuracy of 99% on the detection and classification tasks. Moreover, the study by Maurya et al. used an external validation data set composed of 217 images.

Maron et al.[36] compared the performance of the state of art algorithm to the 112 dermatologists in diagnosing dermatological lesions including BCC. The sensitivity was similar for both (73.8%). However, in terms of specificity, the proposed CNN model scored higher (99.5%) outperforming the dermatologist specificity (97.8%).

BCC detection in OCT images

Two studies[44,45] reported on BCC detection in OCT images. Different types of OCT devices were used for imaging: polarization-sensitive OCT and conventional OCT. The number of images used to train and validate was <100 scans in both studies.

The study by Jorgensen et al. used the approach of the leave-one-out validation strategy to use the whole data set to train and validate the model. In addition, SVM, ensemble and N-tuple classifiers were used to discriminate between BCC and actinic keratosis. Marvdashti et al. used a machine learning algorithm called minimal-redundancy-maximal-relevance to guide the feature selection from multiple features extracted from polarization and the intensity images.

TABLE 1 Articles, models and performance metrics.

Study	Objective	Data set size	Model	Results[b]	External validation
Dermoscopy					
Al-masni et al.[30]	Multiclass skin cancer classification	11,720	Segmentation: FrCN Classification: ResNet-50	Sens = 76%[a] Spec = 98%[a]	No
Igbal et al.[31]	Multiclass skin cancer classification	ISIC-18: 10,015 ISIC-19: 25,331	DCNN	ISIC-18: Sens = 89%[a] Spec = 95%[a] AUC = 0.99 ISIC-19: Sens = 89%[a] Spec = 97%[a] AUC = 0.99	No
Kefel et al.[32]	Binary classifier (BCC vs No BCC)	1378	Logistic regression	AUC = 0.88	No
Khan et al.[33]	Multiclass skin cancer classification	HAM-10,000: 10,015 ISIC-18: 10,015 ISIC-19: 25,331	DenseNet201, MobileNetV2 and ScMFO, Multiclass extreme learning machine	HAM-10,000: Sens = 100% ISIC-18: Sens = 88% ISIC-19: Sens = 85%	No
Kharazmi et al.[34]	Segmentation of cutaneous vasculature in dermoscopy to detect BCC	659	K-means clustering & Random Forest	AUC = 0.97	No
Kharazmi et al.[35]	Binary classifier (BCC vs No BCC)	1199	Sparse auto-encoder	AUC = 0.91	No
Maron et al.[36]	Five-way classification	11,444	ResNet-50	Sens = 74% Spec = 100%	No
Maurya et al.[37]	Segmentation of telangiectasia to detect BCC	630	U-Net	Acc = 99%	Yes
Molina-Molina et al.[38]	Eight-way classification	25,331	Ensemble classifier containing KNN, SVM, linear Gaussian kernels	Sens = 68% Spec = 93%	No
Shimizu et al.[39]	Four-way classification	964	Linear classifier	Detection rate = 83%	No
Wang et al.[40]	Four-way classification	7192	Pretrained CNN from GoogLeNet Inception v3	Sens = 80% Spec = 100%	No
Zhu et al.[41]	Segmentation and classification of 14 skin conditions	13,603	Pretrained EfficientNet-b4 on ImageNet	Sens = 97% Spec = 98%	No
OCT					
Jørgensen et al.[45]	Binary classifier BCC vs actinic keratosis on conventional OCT	76	Linear discriminator, N-tuple classifier, SVM	Acc = 81%	No
Marvdashti et al.[44]	BCC detection on PS-OCT	42	SVM	Acc = 95% Sens = 95% Spec = 95% AUC = 0.97	No
RCM					
Campanella et al.[46]	BCC detection on RCM	365	CNN	AUC = 0.86	Yes

Abbreviations: Acc, accuracy; ANN, artificial neural network; AUC, area under the curve; CNN, convolutional neural network; DCNN, deep convolutional neural network; ELM, Epiluminescence microscopy; FFNN, feedforward neural network; FrCN, full resolution convolutional network; HAM, Human Against Machine, HD-OCT, high-definition OCT; ISIC, International Skin Imaging Collaboration; OCT, optical coherence tomography; PS-OCT, polarization-sensitive optical coherence tomography; RCM, reflectance confocal microscopy; Sens, sensitivity; Spec, specificity; SVM, support vector machine.

[a]Diagnostic estimated calculated from confusion matrix.

[b]Diagnostic estimates were rounded.

Marvdashti et al. reported the model's performance using accuracy (95%), AUC (0.97), sensitivity (95%) and specificity (95%). The accuracy was higher than the accuracy reported by Jorgensen et al. who reported accuracy (81%) as the only primary outcome (Table 1).

BCC detection in RCM images

Campanella et al.[46] investigated the automated detection of BCC in RCM images using a deep-learning model. The internal data set consisted of retrospectively collected 276 stacks of

RCM images from 52 patients after the expert RCM readers excluded stacks that contain malignancies other than BCC and stacks with bad quality. The proposed CNN model was composed of a previously trained ResNet34 backbone along two extra residual block layers and was trained using fivefold cross-validation. In addition, the trained model was further validated on an external data set consisting of 53 stacks. The proposed classifier achieved an AUC of 0.90, 0.88 and 0.86 at the stack level, lesion level and external validation, respectively.

CHALLENGES AND FUTURE DIRECTIONS

In this systematic review, we focussed on the current handcrafted radiomics and deep-learning models for detecting and classification of BCC following different non-invasive imaging techniques. The results of this systematic review indicate a high potential for using AI to improve the diagnostic accuracy for diagnosing BCC by non-invasive diagnostic modalities. Despite the reported potential of AI in medical imaging analysis, the number of studies including radiomics and deep-learning models for detecting BCC is still limited.

The performance of AI models varied substantially between the studies included in the review. Generalizability of AI models depends strongly on the quality and variety of images used both in training and later in testing.[47] However, we attribute the main cause of the differences to the fact that there is significant variation in the models used and amount of data. Unfortunately, the heterogeneity of the data sets acquired for training and the variety in reporting metrics made comparison between the models difficult.

In order to make comparison easier, Luo et al.[48] recommended that future studies on models focussing on classification report their results in specificity, sensitivity, positive and negative predictive values, as well as AUC. In case of predictive models with more than one outcome, accuracy should be additionally reported.

Several challenges currently limit the applicability of the majority of developed algorithms in clinical settings. The major challenge is the lack of repeatability and reproducibility of features. Repeatable features can be defined as features that continue to be the same when imaged multiple times in the same subject with the same acquisition parameters.[49] Reproducible features, on the contrary, are those that remain constant even with a change in the acquisition parameter or the imaged subject.[50]

Repeatability and reproducibility of handcrafted radiomic features (HRFs) have been investigated in several studies. For example, a test–retest study is a well-established method to investigate whether imaging features are reproducible and repeatable.[51] In a test–retest study, medical images are acquired repetitively on the same patient/lesion with exactly the same imaging parameters. A significant number of HRFs appeared not to be reproducible in test–retest scenarios.[52-54] The majority of HRFs appeared to be affected by the changes in imaging parameters indicating that reproducibility of HRFs is also strongly dependent on the variations in these imaging parameters.[55-58] Including reproducibility assessment of HRFs is therefore a necessity for the development of high-quality prediction models.[59]

For deep learning (DL), currently, three major challenges are facing the field. First is the interpretability of DL-based algorithms. Because of their nontransparent nature of providing details about what made the algorithm predict a specific class, DL is considered a black box, which makes it hard to understand and interpret the outcomes.[60] The second challenge is the need for data. In order to train a DL model, enormous amounts of correctly labelled data should be fed to the model.[61] Open access to research data and other digital research resources according to the FAIR principles, assuring the findability, accessibility, interoperability and reuse of data is essential.[62] For dermoscopy, the international imaging initiatives Human Against Machine (HAM) and International Skin Imaging Collaboration (ISIC) are good examples of collective open-source databases, providing many dermoscopy images. Acquiring a large data set of relatively new imaging modalities such as OCT is still challenging, as OCT is not widely used yet. The third challenge is the harmonization of scans acquired with different imaging parameters as heterogeneity in imaging standards could affect the generalizability of the proposed model when used on different data. Harmonization can take place at imaging level or at feature level. At imaging level, harmonization can be achieved through a standardized imaging protocol, whereas at feature level, harmonization could be achieved for example by selecting only high reliable features.[63-66]

There are further challenges that hinder the adoption of AI algorithms in daily dermatological practices. First, privacy concerns, such as sharing dermatological images, which could lead to identifying the patient. Second, there is a lack of standardization in storing and sharing images in dermatology as there is an absence of a unified system such as the picture archiving and communication system used in the field of radiology. Third, the over-representation of some skin conditions and lack of skin tone diversities will lead to the development of biased models that would not be applicable to other conditions or demographics.[67]

CONCLUSION

The reviewed publications demonstrate the potential benefit of AI in the detection of BCC in dermoscopy, OCT and RCM. If the current challenges are overcome, AI tools can be used to develop clinical decision support systems for BCC diagnosis which potentially save time and costs and lead to a higher diagnostic accuracy and safety for the patient.

ACKNOWLEDGEMENTS

We would like to thank Prof. Dr. Mariano Suppa and Dr. Clément Lenoir from Hôpital Erasme Brussels for providing the reflectance confocal microscopy image.

FUNDING INFORMATION
None.

CONFLICT OF INTEREST STATEMENT
Prof. Dr. Lambin is the co-inventor of two nonissued, nonlicensed patents on Deep Learning-Radiomics-Histomics (N2024482, N2024889). Dr. Woodruff and Prof. Dr. Lambin have (minority) shares in the company Radiomics. The rest of the authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. All authors have approved the manuscript and agreed with its submission to the Journal of the European Academy of Dermatology and Venereology.

DATA AVAILABILITY STATEMENT
We confirm that the data supporting the findings of this systematic review are available within the selected articles.

ORCID
Y. Widaatalla https://orcid.org/0000-0002-0652-8661
T. Wolswijk https://orcid.org/0000-0002-0253-8969
F. Adan https://orcid.org/0000-0003-3331-6610
L. M. Hillen https://orcid.org/0000-0002-8808-4952
H. C. Woodruff https://orcid.org/0000-0001-7911-5123
I. Halilaj https://orcid.org/0000-0002-7805-7918
A. Ibrahim https://orcid.org/0000-0003-4138-5755
P. Lambin https://orcid.org/0000-0001-7961-0191
K. Mosterd https://orcid.org/0000-0002-9065-3050

REFERENCES
1. Bichakjian C, Armstrong A, Baum C, Bordeaux JS, Brown M, Busam KJ, et al. Guidelines of care for the management of basal cell carcinoma. J Am Acad Dermatol. 2018;78(3):540–59.
2. Lomas A, Leonardi-Bee J, Bath-Hextall F. A systematic review of worldwide incidence of nonmelanoma skin cancer. Br J Dermatol. 2012;166(5):1069–80.
3. Rogers HW, Weinstock MA, Feldman SR, Coldiron BM. Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the US population, 2012. JAMA Dermatol. 2015;151(10):1081–6.
4. Borgonjen R, van Everdingen J, Bruijnzeel-Koomen C, van de Kerkhof P, Spuls PI. A national study on adherence to a basal cell carcinoma guideline; development of a tool to assess guideline adherence. Br J Dermatol. 2015;172(4):1008–13.
5. Wahie S, Lawrence CM. Wound complications following diagnostic skin biopsies in dermatology inpatients. Arch Dermatol. 2007;143(10):1267–71.
6. Scatliff JH, Morris PJ. From roentgen to magnetic resonance imaging: the history of medical imaging. N C Med J. 2014;75(2):111–3.
7. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, Van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer. 2012;48(4):441–6.
8. Refaee T, Wu G, Ibrahim A, Halilaj I, Leijenaar R, Rogers W, et al. The emerging role of Radiomics in COPD and lung cancer. Respiration; Int Rev Thoracic Dis. 2020;99(2):99–107.
9. Mayerhoefer ME, Materka A, Langs G, Häggström I, Szczypiński P, Gibbs P, et al. Introduction to radiomics. J Nucl Med. 2020;61(4):488–95.
10. Namatēvs I. Deep convolutional neural networks: structure, feature extraction and training. Inform Technol Manage Sci. 2017;20(1):40–7.
11. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology. 2016;278(2):563–77.
12. Najafabadi MM, Villanustre F, Khoshgoftaar TM, Seliya N, Wald R, Muharemagic E. Deep learning applications and challenges in big data analytics. J Big Data. 2015;2(1):1–21.
13. Carli P, De Giorgi V, Argenziano G, Palli D, Giannotti B. Preoperative diagnosis of pigmented skin lesions: in vivo dermoscopy performs better than dermoscopy on photographic images. J Eur Acad Dermatol Venereol. 2002;16(4):339–46.
14. Korotkov K, Garcia R. Computerized analysis of pigmented skin lesions: a review. Artif Intell Med. 2012;56(2):69–90.
15. Braun R, French L, Saurat J. Dermoscopy of pigmented lesions: a valuable tool in the diagnosis of melanoma. Swiss Med Wkly. 2004;134(7–8):83–90.
16. Reiter O, Mimouni I, Gdalevich M, Marghoob AA, Levi A, Hodak E, et al. The diagnostic accuracy of dermoscopy for basal cell carcinoma: a systematic review and meta-analysis. J Am Acad Dermatol. 2019;80(5):1380–8.
17. Conforti C, Giuffrida R, Vezzoni R, Resende FS, di Meo N, Zalaudek I. Dermoscopy and the experienced clinicians. Int J Dermatol. 2020;59(1):16–22.
18. Celebi ME, Aslandogan YA, Bergstresser PR. Unsupervised border detection of skin lesion images. International conference on information technology: coding and computing (ITCC'05)-volume II. Las Vegas: IEEE; 2005.
19. Welzel J, Lankenau E, Birngruber R, Engelhardt R. Optical coherence tomography of the human skin. J Am Acad Dermatol. 1997;37(6):958–63.
20. Schmitt JM, Xiang S, Yung KM. Speckle in optical coherence tomography. J Biomed Opt. 1999;4(1):95–105.
21. Jelly ET, Zhao Y, Chu KK, Price H, Crose M, Steelman ZA, et al. Deep imaging with 1.3 μm dual-axis optical coherence tomography and an enhanced depth of focus. Biomed Opt Express. 2021;12(12):7689–702.
22. Hussain AA, Themstrup L, Jemec GBE. Optical coherence tomography in the diagnosis of basal cell carcinoma. Arch Dermatol Res. 2015;307(1):1–10.
23. Sinx KA, van Loo E, Tonk EH, Kelleners-Smeets NW, Winnepenninckx VJ, Nelemans PJ, et al. Optical coherence tomography for noninvasive diagnosis and subtyping of basal cell carcinoma: a prospective cohort study. J Invest Dermatol. 2020;140(10):1962–7.
24. Adan F, Nelemans PJ, Essers BA, Brinkhuizen T, Dodemont SR, Kessels JP, et al. Optical coherence tomography versus punch biopsy for diagnosis of basal cell carcinoma: a multicentre, randomised, non-inferiority trial. Lancet Oncol. 2022;23(8):1087–96.
25. Marra DE, Torres A, Schanbacher CF, Gonzalez S. Detection of residual basal cell carcinoma by in vivo confocal microscopy. Dermatol Surg. 2005;31(5):538–41.
26. Rajadhyaksha M, González S, Zavislan JM, Anderson RR, Webb RH. In vivo confocal scanning laser microscopy of human skin II: advances in instrumentation and comparison with histology. J Invest Dermatol. 1999;113(3):293–303.
27. Nori S, Rius-Díaz F, Cuevas J, Goldgeier M, Jaen P, Torres A, et al. Sensitivity and specificity of reflectance-mode confocal microscopy for in vivo diagnosis of basal cell carcinoma: a multicenter study. J Am Acad Dermatol. 2004;51(6):923–30.
28. Hupe M. EndNote X9. J Electronic Resources Med Libraries. 2019;16(3–4):117–9.
29. Moher D, Liberati A, Tetzlaff J, Altman DG, Group* P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med. 2009;151(4):264–9.
30. Al-Masni MA, Kim D-H, Kim T-S. Multiple skin lesions diagnostics via integrated deep convolutional networks for segmentation and classification. Comput Methods Programs Biomed. 2020;190:105351.
31. Iqbal I, Younus M, Walayat K, Kakar MU, Ma J. Automated multiclass classification of skin lesions through deep convolutional neural network with dermoscopic images. Comput Med Imaging Graph. 2021;88:101843.

32. Kefel S, Pelin Kefel S, LeAnder RW, Kaur R, Kasmi R, Mishra NK, et al. Adaptable texture-based segmentation by variance and intensity for automatic detection of semitranslucent and pink blush areas in basal cell carcinoma. Skin Res Technol. 2016;22(4):412–22.
33. Attique Khan M, Sharif M, Akram T, Kadry S, Hsu CH. A two-stream deep neural network-based intelligent system for complex skin cancer types classification. Int J Intell Syst. 2022;37(12):10621–49.
34. Kharazmi P, AlJasser MI, Lui H, Wang ZJ, Lee TK. Automated detection and segmentation of vascular structures of skin lesions seen in Dermoscopy, with an application to basal cell carcinoma classification. IEEE J Biomed Health Inform. 2016;21(6):1675–84.
35. Kharazmi P, Kalia S, Lui H, Wang Z, Lee T. A feature fusion system for basal cell carcinoma detection through data-driven feature learning and patient profile. Skin Res Technol. 2018;24(2):256–64.
36. Maron RC, Weichenthal M, Utikal JS, Hekler A, Berking C, Hauschild A, et al. Systematic outperformance of 112 dermatologists in multiclass skin cancer image classification by convolutional neural networks. Eur J Cancer. 2019;119:57–65.
37. Maurya A, Stanley R, Lama N, Jagannathan S, Saeed D, Swinfard S, et al. A deep learning approach to detect blood vessels in basal cell carcinoma. Skin Res Technol. 2022;28:571–6.
38. Molina-Molina EO, Solorza-Calderón S, Álvarez-Borrego J. Classification of dermoscopy skin lesion color-images using fractal-deep learning features. Appl Sci. [Internet] 2020;10:5954. https://doi.org/10.3390/app10175954
39. Shimizu K, Iyatomi H, Celebi ME, Norton K-A, Tanaka M. Four-class classification of skin lesions with task decomposition strategy. IEEE Trans Biomed Eng. 2014;62(1):274–83.
40. Wang SQ, Zhang XY, Liu J, Tao C, Zhu CY, Shu C, et al. Deep learning-based, computer-aided classifier developed with dermoscopic images shows comparable performance to 164 dermatologists in cutaneous disease diagnosis in the Chinese population. Chin Med J (Engl). 2020;133(17):2027–36.
41. Zhu CY, Wang YK, Chen HP, Gao KL, Shu C, Wang JC, et al. A deep learning based framework for diagnosing multiple skin diseases in a clinical environment. Front Med. 2021;8:626369.
42. Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Scientific Data. 2018;5(1):1–9.
43. Combalia M, Codella NC, Rotemberg V, Helba B, Vilaplana V, Reiter O, et al. Bcn20000: Dermoscopic lesions in the wild. *arXiv preprint arXiv:190802288*. 2019.
44. Marvdashti T, Duan L, Aasi SZ, Tang JY, Bowden AKE. Classification of basal cell carcinoma in human skin using machine learning and quantitative features captured by polarization sensitive optical coherence tomography. Biomed Opt Express. 2016;7(9):3721–35.
45. Jørgensen TM, Tycho A, Mogensen M, Bjerring P, Jemec GB. Machine-learning classification of non-melanoma skin cancers from image features obtained by optical coherence tomography. Skin Res Technol. 2008;14(3):364–9.
46. Campanella G, Navarrete-Dechent C, Liopyris K, Monnier J, Aleissa S, Minhas B, et al. Deep learning for basal cell carcinoma detection for reflectance confocal microscopy. J Invest Dermatol. 2022;142(1):97–103.
47. Willemink MJ, Koszek WA, Hardell C, Wu J, Fleischmann D, Harvey H, et al. Preparing medical imaging data for machine learning. Radiology. 2020;295(1):4–15.
48. Luo W, Phung D, Tran T, Gupta S, Rana S, Karmakar C, et al. Guidelines for developing and reporting machine learning predictive models in biomedical research: a multidisciplinary view. J Med Internet Res. 2016;18(12):e5870.
49. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: the process and the challenges. Magn Reson Imaging. 2012;30(9):1234–48.
50. O'Connor JP, Aboagye EO, Adams JE, Aerts HJ, Barrington SF, Beer AJ, et al. Imaging biomarker roadmap for cancer studies. Nat Rev Clin Oncol. 2017;14(3):169–86.
51. Hiremath A, Shiradkar R, Merisaari H, Prasanna P, Ettala O, Taimen P, et al. Test-retest repeatability of a deep learning architecture in detecting and segmenting clinically significant prostate cancer on apparent diffusion coefficient (ADC) maps. Eur Radiol. 2021;31(1):379–91.
52. Leijenaar RT, Carvalho S, Velazquez ER, Van Elmpt WJ, Parmar C, Hoekstra OS, et al. Stability of FDG-PET radiomics features: an integrated analysis of test-retest and inter-observer variability. Acta Oncol. 2013;52(7):1391–7.
53. van Timmeren JE, Leijenaar RT, van Elmpt W, Wang J, Zhang Z, Dekker A, et al. Test–retest data for radiomics feature stability analysis: generalizable or study-specific? Tomography. 2016;2(4):361–5.
54. Shiri I, Abdollahi H, Shaysteh S, Mahdavi SR. Test–retest reproducibility and robustness analysis of recurrent glioblastoma MRI radiomics texture features. Iranian J Radiol. 2017;14:e48035.
55. Baeßler B, Weiss K, Dos Santos DP. Robustness and reproducibility of radiomics in magnetic resonance imaging: a phantom study. Invest Radiol. 2019;54(4):221–8.
56. Zhao B, Tan Y, Tsai W-Y, Qi J, Xie C, Lu L, et al. Reproducibility of radiomics for deciphering tumor phenotype with imaging. Sci Rep. 2016;6(1):1–7.
57. Ibrahim A, Refaee T, Leijenaar RT, Primakov S, Hustinx R, Mottaghy FM, et al. The application of a workflow integrating the variable reproducibility and harmonizability of radiomic features on a phantom dataset. PLoS One. 2021;16(5):e0251147.
58. Ibrahim A, Refaee T, Primakov S, Barufaldi B, Acciavatti RJ, Granzier RW, et al. The effects of in-plane spatial resolution on CT-based radiomic features' stability with and without ComBat harmonization. Cancer. 2021;13(8):1848.
59. Mali SA, Ibrahim A, Woodruff HC, Andrearczyk V, Müller H, Primakov S, et al. Making radiomics more reproducible across scanner and imaging protocol variations: a review of harmonization methods. J Personalized Med. 2021;11(9):842.
60. Salahuddin Z, Woodruff HC, Chatterjee A, Lambin P. Transparency of deep neural networks for medical image analysis: a review of interpretability methods. Comput Biol Med. 2022;140:105111.
61. Giger ML. Machine learning in medical imaging. J Am Coll Radiol. 2018;15(3):512–20.
62. Yue P, Shangguan B, Hu L, Jiang L, Zhang C, Cao Z, et al. Towards a training data model for artificial intelligence in earth observation. Int J Geograph Inform Sci. 2022;36(11):2113–37.
63. Da-Ano R, Visvikis D, Hatt M. Harmonization strategies for multicenter radiomics investigations. Phys Med Biol. 2020;65(24):24TR02.
64. Ibrahim A, Primakov S, Beuque M, Woodruff H, Halilaj I, Wu G, et al. Radiomics for precision medicine: current challenges, future prospects, and the proposal of a new framework. Methods. 2021;188:20–9.
65. Rogers W, Thulasi Seetha S, Refaee TA, Lieverse RI, Granzier RW, Ibrahim A, et al. Radiomics: from qualitative to quantitative imaging. Br J Radiol. 2020;93(1108):20190948.
66. Papadimitroulas P, Brocki L, Chung NC, Marchadour W, Vermet F, Gaubert L, et al. Artificial intelligence: deep learning in oncological radiomics and challenges of interpretability and data harmonization. Phys Med. 2021;83:108–21.
67. Eapen BR. Artificial intelligence in dermatology: a practical introduction to a paradigm shift. Indian Dermatol Online J. 2020;11(6):881–9.

How to cite this article: Widaatalla Y, Wolswijk T, Adan F, Hillen LM, Woodruff HC, Halilaj I, et al. The application of artificial intelligence in the detection of basal cell carcinoma: A systematic review. J Eur Acad Dermatol Venereol. 2023;00:1–8. https://doi.org/10.1111/jdv.18963