Proceedings of ICBioMed 2025 Symposium: AI for Healthcare: Advanced Medical Data Analytics and Smart Rehabilitation
DOI: 10.54254/2753-8818/2025.AU27459

A Review of Applications of Artificial Intelligence in Skin Tumor Image Recognition Across Different Skin Types

Yu-Ching Chiu

Department of Clinical Medicine, Guangzhou Medical University, Guangzhou, China
chiuyuching17@gmail.com

Abstract. Skin cancer remains one of the most common malignancies worldwide, with an increasing incidence in recent decades. Early detection is essential to reduce mortality, but traditional diagnostic methods, including visual examination and dermoscopy, are time-consuming and susceptible to inter-observer variability. Artificial intelligence (AI), particularly deep learning, has shown great potential in medical image analysis and skin tumor identification, reaching the level of dermatologists. This review systematically analyzes recent advances in AI in image recognition for skin tumors, with a particular focus on the equity and performance of these models among different skin types and ethnic groups. Through a literature review of studies published in recent years, this study examined public datasets, algorithmic approaches, equity challenges, and clinical integration prospects. The review found that while AI systems show robust diagnostic performance, there are still differences in model accuracy across skin tones due to imbalanced datasets. We conclude that enhancing dataset diversity, improving interpretability, and establishing clinical validation pipelines are essential for achieving equitable AI in dermatology.

Keywords: AI, Skin cancer, Image recognition, Skin tumor

1. Introduction

There are many different types of skin cancer, including melanoma, squamous cell carcinoma (SCC), basal cell carcinoma (BCC), etc. [1]. Skin cancer is one of the most common tumors and one of the major public health challenges worldwide. There are more than 1.5 million new cases of skin cancer each year, and the mortality rate varies significantly between regions [2], bringing a huge economic burden on health [3]. Although melanoma has a low incidence, up to 75% of skin cancer deaths are caused by melanoma, which has the potential to metastasize to other parts of the body [4, 5]. The 5-year survival rate for early detection of melanoma is more than 90%, but it declines dramatically in advanced disease [6, 7], so early and accurate detection of skin cancer types is essential to guide appropriate treatment, and management, and improve morbidity, mortality, and prognosis [8].

Traditional diagnostic methods, such as clinical examination and dermatoscopy, are often limited by the expertise and subjectivity of doctors. With the rise of computational methods, artificial intelligence (AI) has emerged as a promising tool, exploding significantly in medical applications. Moreover, with the routine acquisition of images of suspicious skin lesions in dermatology, the field

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2025.AU27459

of skin cancer is expected to utilize new and evolving machine learning techniques [8]. Deep learning-based convolutional neural networks (CNNs) have shown comparable or better performance than dermatologists in classifying skin lesions based on dermatoscopic images [9, 10]. A landmark study by Esteva et al. confirmed that CNNs trained on large clinical image datasets can achieve dermatologist-level accuracy in distinguishing between Keratinocytic carcinoma and melanoma [11]. Recently, Transformer-based models and hybrid architectures have further advanced recognition capabilities [12, 13]. These advances show that deep learning is driving skin tumor identification into a new stage of greater intelligence and precision.

This paper provides a comprehensive summary and analysis of differences in the diagnosis of skin tumors of different skin color types and provides a reference for researchers to identify current deficiencies and bottlenecks. This review not only provides valuable guidance for improving the diagnostic accuracy of patients with dark skin color but also provides new directions for promoting cross-skin color equity, improving dataset construction, and improving model interpretability and clinical usability at the macro level. These discussions will lay the foundation for subsequent academic research and practical exploration and provide a reference for achieving equitable applications of artificial intelligence in dermatology.

2. Public datasets and data preparation

2.1. Importance of skin tumor image data

The development of effective AI models in dermatology relies critically on large-scale, high-quality datasets with precise diagnostic annotations, as the performance of such models is fundamentally determined by the quantity, quality, and accuracy of the training images, which must encompass both dermoscopic and clinical photographs [11,14]. However, most data have limited representation of darker skin tones, resulting in systemic bias, poor diagnostic performance, and inadequate generalization power [15-18]. Therefore, it is crucial to maintain quality control of skin tumor image data. Data enhancement techniques, including rotation, flipping, and illumination adjustment, can be used to alleviate imbalances in datasets.

2.2. Commonly used datasets

Several publicly available datasets have accelerated research on skin lesion identification. Fitzpatrick17k is the first large-scale dataset to explicitly provide skin color labeling for fairness research**. ISIC Archive ISIC Archive/HAM10000 (International Collaborative on Skin Imaging) offers the largest collection of open-access dermatoscopic images, widely used in artificial intelligence competitions [19]. Other datasets, such as PH2 and Dermofit, provide smaller but highly selected samples with expert annotations [20]. These datasets typically include images of melanocytic nevus, melanoma, and keratinocytic carcinoma, with varying resolution and quality. However, it is important to note that there is a significant imbalance in skin color diversity in the existing public databases. One of the main issues mentioned in several articles is the use of non-diverse datasets to train CNNs [21, 22], where most samples are currently from light-skinned European and Asian populations, while dark-skinned patients tend to have less than 10% [17,23]. Only Fitzpatrick17k [15] and DDI [24] made breakthroughs in skin color labeling and equilibrium. This distribution bias may lead to the model's excellent performance in the white population (AUC can reach above 0.90), but the sensitivity and specificity are significantly reduced in dark-skinned or

DOI: 10.54254/2753-8818/2025.AU27459

ethnic minority groups [17]. Therefore, the representativeness of the database directly affects the fairness and generalization ability of AI models among different populations.

2.3. Data annotation and preprocessing

Annotation remains a bottleneck because mislabeling or inconsistent classification can lead to biased models [18]. To improve model performance, preprocessing steps such as image cropping, normalization, enhancement (rotation, flipping, and scaling), and artifact removal are commonly applied [25]. Data augmentation not only increases effective dataset size but also improves model robustness by simulating real-world variability.

In the context of different skin tones, the challenges of data labeling and preprocessing are even more prominent. Dark skin color often reduces the contrast between the lesion and the surrounding skin, increases the difficulty of labeling, and is prone to misclassification [17]. To alleviate this problem, researchers have tried to introduce skin-tone-related lighting and color transformations in data enhancement, such as artificially "darkened" images to train datasets to simulate clinical images under multiple skin tones, thereby improving the robustness and fairness of the model in diverse populations [26,27]. In addition, constructing expert consistency labeling standards across skin tones is also a key part of ensuring the fairness of AI models [24].

3. Artificial intelligence algorithms

3.1. Traditional machine learning methods

Traditional machine learning methods, such as support vector machines (SVMs) and random forests, rely on hand-crafted features, limiting their performance in complex lesion identification. While these models have had some success, they require domain expertise to design relevant features and have limited universality. For example, early studies used color histograms, texture features (e.g., GLCM grayscale symbiosis matrix), and marginal information to distinguish melanoma from moles [28]. However, these features performed better in light-skinned samples, but in darker skin tones, the sensitivity and specificity of the model decreased significantly due to the low contrast between the lesion and the skin background [17]. Therefore, traditional machine learning methods are particularly limited in generalization capabilities across skin colors.

3.2. Deep learning methods

Deep learning, particularly CNN architectures such as AlexNet, VGG, ResNet, and Inception, has demonstrated state-of-the-art performance in dermatoscopic image classification. The introduction of CNNs has revolutionized the identification of skin lesions. Architectures such as AlexNet, VGG, ResNet, and Inception have been widely used to achieve dermatologist-grade accuracy in the classification of malignancies such as melanoma [11,29]. Transformer-based models (e.g., Vision Transformer, Swin Transformer) and transfer learning further improve accuracy and generalizability. One study showed that Swin Transformers demonstrated superior scalability in medical imaging tasks[12]. These advances have significantly improved tumor identification performance, but there are still significant differences among people of different skin tones. For example, the AUC of CNNs often reaches 0.93–0.95 on light skin tones and decreases to 0.84–0.87 in dark skin tones [17].

Relying on automated feature extraction and multimodal information fusion, deep learning significantly improves the accuracy and intelligence of tumor identification in medical images. Its

DOI: 10.54254/2753-8818/2025.AU27459

application not only promotes early screening and precise segmentation of tumors, but also demonstrates capabilities in personalized diagnosis, treatment, and prognosis prediction.

4. AI in medical imaging for tumor detection across skin tones

Recent studies have shown that the effectiveness of AI models in people with different skin tones varies significantly. Table 1 summarizes recent studies evaluating dermatology AI performance across different Fitzpatrick skin tones. While artificial intelligence models generally achieved high diagnostic accuracy for light skin tones (FST 1–3), their performance consistently declined in darker skin tones (FST 4–5), with notable reductions in sensitivity, specificity, and segmentation quality.

As the application of artificial intelligence in skin tumor diagnosis continues to expand, its fairness among different skin color groups has gradually attracted attention. One study showed that the overall performance of the three mainstream public models (ModelDerm, DeepDerm, and CNNs trained on HAM10000) was significantly lower than that of the original test set in DDI (Diverse Dermatology Images), a multi-skin skin image database with a nearly balanced skin tone ratio. The AUC of the model in dark-skinned people was only 0.50–0.57, and the sensitivity difference was particularly prominent [17]. This shows that the performance gap persists even under balanced sample sizes. Another study showed a significant decline in the performance of a CNN-based skin lesion classifier when used on dark-skinned individuals, due to biases in its training data [30]. Aggarwal et al. highlighted that most deep learning dermatology studies relied on datasets dominated by light-skinned patients, raising concerns that CNN diagnostic performance may decline significantly in dark-skinned populations due to poor representativeness of training data [31]. Moreover, the results of Barros et al. showed that deep learning models performed best on light skin tones, while melanoma recognition performance on medium and dark skin tones decreased significantly or was even completely unrecognized [32]. Overall, these studies consistently demonstrate that current dermatology AI models underperform in dark-skinned populations, underscoring significant fairness risks. Dark-skinned lesions have lower phenotypic contrast/more complex morphology, models are more difficult to learn, and they require more samples to achieve equivalent performance. These findings highlight the importance of improving algorithm fairness and enhancing diverse training data in the future.

DOI: 10.54254/2753-8818/2025.AU27459

Table 1. Dermatology AI performance by skin tone

Ref	Data/AI Model	Performance in Light Skin (FST 1–3)	Performance in Dark Skin (FST 4–5)	Fairness Challenge
Daneshjou et al., [17]	Diverse,curated clinical image set (7,522 images, 64 diseases); tested multiple AI classifiers	AUC ≈ 0.90 in light skin	AUC dropped to 0.74–0.80 in dark skin	Clear disparity: same model significantly less accurate on dark skin, risking underdiagnosis
Benčević et al., [30]	Deeplearning segmentation (U-Net, DeepLab) on skin lesion datasets	Dice score up to 0.85 (lesion segmentation accurate on light skin)	Dice score fell to 0.68–0.72 (lesion boundary errors frequent in dark skin)	Segmentation bias: models less reliable on darker tones due to low contrast and dataset imbalance
Aggarwal, [26]	Fitzpatrick17k dataset (16,577 images, CNN-based models)	Sensitivity ~89–91%, Specificity ~87% for FST 1–3	Sensitivity dropped to ~74%, Specificity ~70–72% for FST 4–5	Models underperform on higher Fitzpatrick classes; emphasized systematic bias and risk of inequitable care
Barros et al., [32]	CNNs trained on Derm7pt + ISIC vs tested on Black skin datasets (acral, nail, mucosal melanoma)	Not primary focus; light skin accuracy reported ~85–88%	Accuracy fell below 65%; melanoma often missed in acral/mucosal sites	Models underperform on higher Fitzpatrick classes; highlighted urgent need for representative datasets

To solve this problem, the researchers proposed a variety of improvements: One is to increase the proportion of dark skin samples in the training dataset to mitigate the bias. Daneshjou R et al created a diverse dermatology image (DDI) dataset on which AI models fine-tuned can close the performance gap between light and dark skin tones [17]. Different learning methods can be used to improve the performance of the model in minority groups as well. For example, Pakzad et al. proposed color-invariant representation learning (CIRCLe), in which the active constraint model learns "skin color invariant features," significantly reducing skin color deviation while maintaining high diagnostic performance [33]. Du et al. proposed disentangled contrastive learning, which effectively eliminates the interference of skin color information and realizes "consistent diagnosis and skin color irrelevance" learning [34]. In addition, human-robot collaboration [35] can also make up for the shortcomings of AI in special skin color recognition.

5. Discussion: fairness and clinical integration

5.1. The promising role of AI in democratizing dermatology

In current dermatology workflows, there are differences in the diagnosis and outcomes of skin cancers, with patients with skin of color being diagnosed at a later stage, leading to increased morbidity, mortality, and cost [28,36]. Therefore, introducing AI in the clinic to assist in the recognition of skin tumor images as a supportive diagnostic tool can improve diagnostic consistency, reduce delays, and promote equitable healthcare.

In controlled studies, AI systems have shown comparable performance to dermatologists [10,35]. In real-world applications, including mobile-based diagnostic applications, they enable patients or primary care physicians to capture images and undergo AI-driven risk assessments [37]. Such systems are particularly valuable in settings with limited resources and limited dermatologist visits [38]. These findings underscore the potential of AI as a supportive diagnostic tool.

DOI: 10.54254/2753-8818/2025.AU27459

5.2. Challenges and progress in AI for skin cancer detection across skin colors

At present, the performance of artificial intelligence models among different skin colors and ethnicities is not universal, and there are many serious flaws in the existing clinical evaluation. This can perpetuate adverse health outcomes in communities of color [39]. Clinical practice also shows that although mobile AI applications [40,41] have potential, they require stricter regulation and evaluation in different skin color patient groups. For example, AI algorithms used highly suspicious skin images that had been selected by dermatologists for biopsy for development and evaluation, and the data results may be biased, increasing the sensitivity and specificity of the results. In addition, the variability in the quality of image acquisition can affect the performance of the algorithm, resulting in images that cannot be evaluated, which seriously affects usability in practice [40]. Similarly, for rare diagnoses and atypical tumor manifestations, the accuracy of diagnostic results is significantly reduced due to the small number of images used for training [42]. Data privacy remains a critical concern as well, especially when deploying cloud-based diagnostic systems [43].

Regulatory frameworks such as the FDA approval process are still evolving for AI-driven diagnostics. Currently, the application of artificial intelligence for the detection of skin cancer is rarely EU CE marked, and very few have been approved by the FDA, limiting the widespread clinical use of products under development [42]. MelaFind, which was approved by the FDA in 2011, was discontinued or used clinically in 2017 due to its low specificity, which led to many unnecessary biopsies. It was only used for pigmented lesions, and could not adapt to clinical workflows and reimbursement environments [44]. Furthermore, the mechanism of action of AI has been described as a "black box" because the algorithms themselves and the training sets they are based on cannot be accessed by physicians using them for evaluation, which also raises concerns about transparency and accountability and the need for ongoing research into explainable artificial intelligence (XAI) [45, 46].

To fully realise the value of artificial intelligence in skin tumour diagnosis, it is critical to both optimize its performance and ensure diagnostic accuracy across different skin tones. This requires dermatology datasets that include representative images of skin of color during both training and testing. Moreover, researchers should progress the theoretical and methodological foundations of AI, with an emphasis on algorithmic fairness and interpretability.

6. Conclusion

This study aims to investigate the diagnostic performance of artificial intelligence (AI) across different skin tones in dermatology. Findings indicate that existing AI models have yet to achieve universal applicability and comprehensive adoption. AI models typically perform with high diagnostic accuracy in fair-skinned populations, but diagnostic accuracy and sensitivity both decline significantly in darker-skinned populations. Nevertheless, the study also identified efforts by some scientists to address this issue by increasing the proportion of dark-skinned samples or developing novel learning strategies. This research provides comprehensive and summarized evidence of the directions that require improvement in AI's application to tumor image recognition, which is important for both the development of machine learning in AI and the diagnosis of skin tumors. A potential limitation of this study is that it contains a limited number of references, and most of them are theoretical studies that have not been applied in real clinical practice, resulting in insufficient evidence levels which may affect the generalizability of the results. Therefore, future research should increase the sample size and focus on clinical practice validation. Overall, this study provides

DOI: 10.54254/2753-8818/2025.AU27459

new insights into the development of AI in dermatology and emphasizes the importance of improving fairness across skin tones and clinical integration.

References

[1] Zhao, Y., et al., Advances in the Relationship Between Skin Tumor Occurrence, Development, Prognosis, and the Human Microbiome. Curr Microbiol, 2025. 82(9): p. 408.
[2] Lomas, A., J. Leonardi-Bee, and F. Bath-Hextall, A systematic review of worldwide incidence of nonmelanoma skin cancer. Br J Dermatol, 2012. 166(5): p. 1069-80.
[3] Zink, A., Trends in the treatment and prevention of keratinocyte carcinoma (non-melanoma skin cancer). Curr Opin Pharmacol, 2019. 46: p. 19-23.
[4] Siegel, R.L., K.D. Miller, and A. Jemal, Cancer statistics, 2020. CA Cancer J Clin, 2020. 70(1): p. 7-30.
[5] statistics, C.R.U.S.c.i.; Available from: https: //www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/skin-cancer/incidence2017.
[6] Karimkhani, C., et al., The global burden of melanoma: results from the Global Burden of Disease Study 2015. Br J Dermatol, 2017. 177(1): p. 134-140.
[7] Balch, C.M., et al., Final version of 2009 AJCC melanoma staging and classification. J Clin Oncol, 2009. 27(36): p. 6199-206.
[8] Ferrante di Ruffano, L., et al., Computer-assisted diagnosis techniques (dermoscopy and spectroscopy-based) for diagnosing skin cancer in adults. Cochrane Database Syst Rev, 2018. 12(12): p. Cd013186.
[9] Brinker, T.J., et al., A convolutional neural network trained with dermoscopic images performed on par with 145 dermatologists in a clinical melanoma image classification task. Eur J Cancer, 2019. 111: p. 148-154.
[10] Haenssle, H.A., et al., Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Ann Oncol, 2018. 29(8): p. 1836-1842.
[11] Esteva, A., et al., Dermatologist-level classification of skin cancer with deep neural networks. Nature, 2017. 542(7639): p. 115-118.
[12] Liu, Z., et al. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. in 2021 IEEE/CVF International Conference on Computer Vision (ICCV). 2021.
[13] Dosovitskiy, A., et al., An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. 2021.
[14] Sanchez, K., et al., Clinical Application of Artificial Intelligence for Non-melanoma Skin Cancer. Curr Treat Options Oncol, 2023. 24(4): p. 373-379.
[15] Groh, M., et al. Evaluating Deep Neural Networks Trained on Clinical Images in Dermatology with the Fitzpatrick 17k Dataset. in 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2021.
[16] Adamson, A.S. and A. Smith, Machine Learning and Health Care Disparities in Dermatology. JAMA Dermatol, 2018. 154(11): p. 1247-1248.
[17] Daneshjou, R., et al., Disparities in dermatology AI performance on a diverse, curated clinical image set. Sci Adv, 2022. 8(32): p. eabq6147.
[18] Pacheco, A.G.C. and R.A. Krohling, Recent advances in deep learning applied to skin cancer detection. ArXiv, 2019. abs/1912.03280.
[19] Codella, N., et al., Skin Lesion Analysis Toward Melanoma Detection 2018: A Challenge Hosted by the International Skin Imaging Collaboration (ISIC). 2019.
[20] Mendonca, T., et al., PH² - a dermoscopic image database for research and benchmarking. Annu Int Conf IEEE Eng Med Biol Soc, 2013. 2013: p. 5437-40.
[21] Kamulegeya, L., et al., Using artificial intelligence on dermatology conditions in Uganda: a case for diversity in training data sets for machine learning. Afr Health Sci, 2023. 23(2): p. 753-763.
[22] Eapen, B.R., Artificial Intelligence in Dermatology: A Practical Introduction to a Paradigm Shift. Indian Dermatol Online J, 2020. 11(6): p. 881-889.
[23] Trager, M.H., et al., Artificial intelligence for nonmelanoma skin cancer. Clin Dermatol, 2024. 42(5): p. 466-476.
[24] Chiu, M.C., et al. DDI-CoCo: A Dataset for Understanding the Effect of Color Contrast in Machine-Assisted Skin Disease Detection. in ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2024.
[25] Shorten, C. and T.M. Khoshgoftaar, A survey on Image Data Augmentation for Deep Learning. Journal of Big Data, 2019. 6: p. 1-48.

DOI: 10.54254/2753-8818/2025.AU27459

[26] Aggarwal, P. and F.A. Papay, Artificial intelligence image recognition of melanoma and basal cell carcinoma in racially diverse populations. J Dermatolog Treat, 2022. 33(4): p. 2257-2262.
[27] Kumar, A., G. Hamarneh, and M.S. Drew. Illumination-based Transformations Improve Skin Lesion Segmentation in Dermoscopic Images. in 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2020.
[28] Barata, C., M.E. Celebi, and J.S. Marques, A Survey of Feature Extraction in Dermoscopy Image Analysis of Skin Cancer. IEEE Journal of Biomedical and Health Informatics, 2019. 23(3): p. 1096-1109.
[29] Kawahara, J., A. Bentaieb, and G. Hamarneh, Deep features to classify skin lesions. 2016. 1397-1400.
[30] Benčević, M., et al., Understanding skin color bias in deep learning-based skin lesion segmentation. Comput Methods Programs Biomed, 2024. 245: p. 108044.
[31] Aggarwal, P., Performance of Artificial Intelligence Imaging Models in Detecting Dermatological Manifestations in Higher Fitzpatrick Skin Color Classifications. JMIR Dermatol, 2021. 4(2): p. e31697.
[32] Barros, L., L. Chaves, and S. Avila. Assessing the generalizability of deep neural networks-based models for black skin lesions. in Iberoamerican Congress on Pattern Recognition. 2023. Springer.
[33] Pakzad, A., K. Abhishek, and G. Hamarneh. CIRCLe: Color invariant representation learning for unbiased classification of skin lesions. in European Conference on Computer Vision. 2022. Springer.
[34] Du, S., et al. Fairdisco: Fairer ai in dermatology via disentanglement contrastive learning. in European Conference on Computer Vision. 2022. Springer.
[35] Tschandl, P., et al., Human-computer collaboration for skin cancer recognition. Nat Med, 2020. 26(8): p. 1229-1234.
[36] Tajbakhsh, N., et al., Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? IEEE Transactions on Medical Imaging, 2016. 35(5): p. 1299-1312.
[37] Markun, S., et al., Mobile teledermatology for skin cancer screening: A diagnostic accuracy study. Medicine (Baltimore), 2017. 96(10): p. e6278.
[38] Kassianos, A.P., et al., Smartphone applications for melanoma detection by community, patient and generalist clinician users: a review. Br J Dermatol, 2015. 172(6): p. 1507-1518.
[39] Guo, L.N., et al., Bias in, bias out: Underreporting and underrepresentation of diverse skin types in machine learning research for skin cancer detection-A scoping review. J Am Acad Dermatol, 2022. 87(1): p. 157-159.
[40] Matin, R.N. and J. Dinnes, AI-based smartphone apps for risk assessment of skin cancer need more evaluation and better regulation. Br J Cancer, 2021. 124(11): p. 1749-1750.
[41] Wongvibulsin, S., et al., Current State of Dermatology Mobile Applications With Artificial Intelligence Features. JAMA Dermatol, 2024. 160(6): p. 646-650.
[42] Beltrami, E.J., et al., Artificial intelligence in the detection of skin cancer. J Am Acad Dermatol, 2022. 87(6): p. 1336-1342.
[43] Price, W.N., 2nd and I.G. Cohen, Privacy in the age of medical big data. Nat Med, 2019. 25(1): p. 37-43.
[44] Times, T.N.Y. Dissent over a device to help find melanoma. 2013 30th Apr 2022]; Available from: https://www.nytimes.com/2013/07/21/business/dissent-over-a-device-to-help-find-melanoma.html.
[45] Holzinger, A., et al., Causability and explainability of artificial intelligence in medicine. Wiley Interdiscip Rev Data Min Knowl Discov, 2019. 9(4): p. e1312.
[46] Hauser, K., et al., Explainable artificial intelligence in skin cancer recognition: A systematic review. Eur J Cancer, 2022. 167: p. 54-69.