Review

Machine Learning Methods in Skin Disease Recognition: A Systematic Review

Jie Sun, Kai Yao, Guangyao Huang, Chengrui Zhang, Mark Leach, Kaizhu Huang and Xi Yang

Special Issue

Machine Learning in Biomaterials, Biostructures and Bioinformatics

Edited by
Dr. Jie Sun, Prof. Dr. Kaizhu Huang and Dr. Sanjairaj Vijayavenkataraman

Review

Machine Learning Methods in Skin Disease Recognition: A Systematic Review

Jie Sun [1,*], Kai Yao [1,2], Guangyao Huang [1], Chengrui Zhang [1,2], Mark Leach [1,*], Kaizhu Huang [3] and Xi Yang [1]

1   School of Advanced Technology, Xi'an Jiaotong-Liverpool University, Suzhou 215123, China
2   School of Engineering, University of Liverpool, Liverpool L69 3BX, UK
3   Data Science Research Centre, Duke Kunshan University, Kunshan 215316, China
*   Correspondence: jie.sun@xjtlu.edu.cn (J.S.); mark.leach@xjtlu.edu.cn (M.L.)

Abstract: Skin lesions affect millions of people worldwide. They can be easily recognized based on their typically abnormal texture and color but are difficult to diagnose due to similar symptoms among certain types of lesions. The motivation for this study is to collate and analyze machine learning (ML) applications in skin lesion research, with the goal of encouraging the development of automated systems for skin disease diagnosis. To assist dermatologists in their clinical diagnosis, several skin image datasets have been developed and published online. Such efforts have motivated researchers and medical staff to develop automatic skin diagnosis systems using image segmentation and classification processes. This paper summarizes the fundamental steps in skin lesion diagnosis based on papers mainly published since 2013. The applications of ML methods (including traditional ML and deep learning (DL)) in skin disease recognition are reviewed based on their contributions, methods, and achieved results. Such technical analysis is beneficial to the continuing development of reliable and effective computer-aided skin disease diagnosis systems. We believe that more research efforts will lead to the current automatic skin diagnosis studies being used in real clinical settings in the near future.

Keywords: skin image segmentation; skin lesion classification; machine learning; deep learning; computer assisted diagnostics; dermatology

1. Introduction

The skin protects the human body from outside hazardous substances and is the largest organ in the human body. Skin diseases can be induced by various causes, such as fungal infections, bacterial infections, allergies, or viruses [1]. Skin lesions, as one of the most common diseases, have affected a large population worldwide.

A skin lesion can be easily recognized based on its abnormal texture and/or color, but it can be difficult to accurately diagnose due to similar symptoms occurring with different types of lesions. Figure 1a,b, show that both contact dermatitis and eczema present similarly with redness, swelling, and chapping on a hand. Similarly, as seen in Figure 1c,d, measles and keratosis show red dots that are sporadically distributed on the skin. These similarities lower diagnosis accuracy, particularly for dermatologists with less than two years of clinical experience. To administer appropriate treatment, it is essential to identify skin lesions correctly and as early as possible. Early diagnosis usually leads to a better prognosis and increases the chance of full recovery in most instances. The training of a dermatologist requires several years of clinical experience as well as high education costs. The average tuition fee in medical school was USD 218,792 in 2021 has risen to around USD 1500 each year since 2013 [2]. An automated solution for skin disease diagnosis is becoming increasingly urgent. This is especially true in developing countries, which often lack the advanced medical equipment and expertise
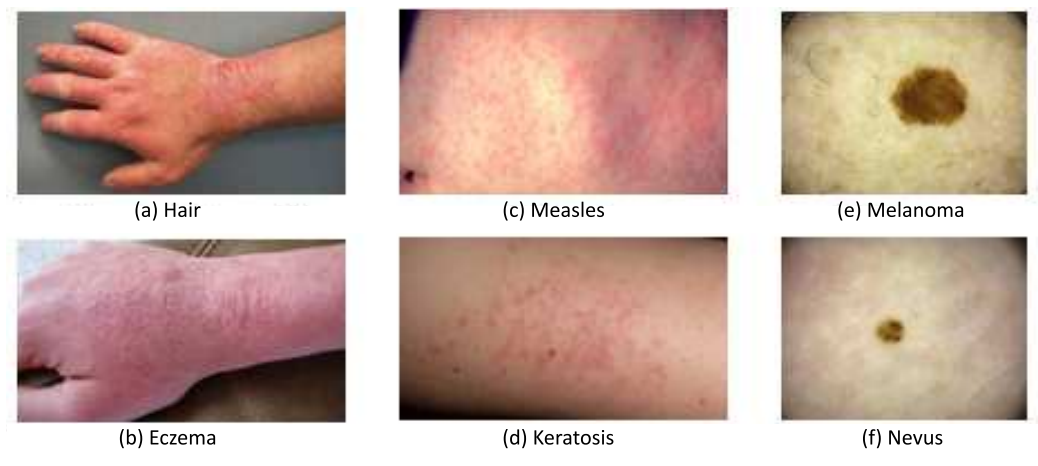
Figure 1. Common skin diseases [2,3].

Artificial intelligence (AI) has made rapid progress in image recognition over the last decade, using a combination of methods such as machine learning (ML) and deep learning (DL). Segmentation and classification models can be developed using traditional ML and DL methods. Traditional ML methods have simpler structures compared to DL methods, which use extracted and selected features as inputs [4]. The DL method can automatically discover underlying patterns and identify the most descriptive and salient features in image recognition and processing tasks. This progress has significantly motivated medical workers to explore the potential for the application of AI methods in disease diagnosis, particularly for skin disease diagnosis. Moreover, DL has already demonstrated its capability in this field by achieving similar diagnostic accuracy as dermatologists with 5 years of clinical experience [5]. Therefore, DL is considered a potential tool for cost-effective skin health diagnosis.

The success of AI applications heavily relies on the support of big data to ensure reliable performance and generalization ability. To speed up the progress of AI application in skin disease diagnosis, it is essential to establish reliable databases. Researchers from multiple organizations worked together to create the International Skin Imaging Collaboration (ISIC) Dataset for skin disease research and organized challenges from 2016 to 2020 [6]. The most common research task is to detect melanoma, a life-threatening skin cancer. According to the World Health Organization (WHO), the number of melanoma cases is expected to increase to 466,914 with associated deaths increasing to 105,904 by 2040 [7].

Several additional datasets with macroscopic or dermoscopic images are publicly accessible on the internet, such as the PH2 dataset [8], the Human Against Machine with 10,000 training images (HAM 10,000) dataset [9], the BCN 20,000 dataset [10], the Interactive Atlas of Dermoscopy (EDRA) dataset [11], and the Med-Node dataset [12]. Such datasets can be used individually, partially, or combined based on the specific tasks under investigation.

Macroscopic or close-up images of skin lesions on the human body are often taken using common digital cameras or smartphone cameras. Dermoscopic images, on the other hand, are collected using a standard procedure in preliminary diagnosis and produce images containing fewer artifacts and more detailed features, which can visualize the skin's inner layer (invisible to the naked eye). Although images may vary with shooting distance, illumination conditions, and camera resolution, dermoscopic images are more likely to achieve higher performance in segmentation and classification than macroscopic images [13].

To conduct a rigorous and thorough review of skin lesion detection studies, we utilized a systematic search strategy that involved the PubMed database—an authoritative, widely-used resource for accessing biomedical and life sciences literature. Our analysis primarily

focused on studies published after 2013 that explored the potential applications of ML in skin lesion diagnosis. To facilitate this analysis, we sourced high-quality datasets from reputable educational and medical organizations. These datasets served as valuable tools not only for developing and testing new diagnostic algorithms but also for educating the broader public on skin health and the importance of early detection. By utilizing these datasets in our research, we gained a deeper understanding of the current state-of-the-art in skin lesion detection and identified areas for future research and innovation.

We discuss the most relevant skin recognition studies, including datasets, image preprocessing, up-to-date traditional ML and DL applications in segmentation and classification, and challenges based on the original research or review papers from journals and scientific conferences in ScienceDirect, IEEE, and SpringerLink databases. The common schematic diagram of the automated skin image diagnosis procedure is shown in Figure 2.
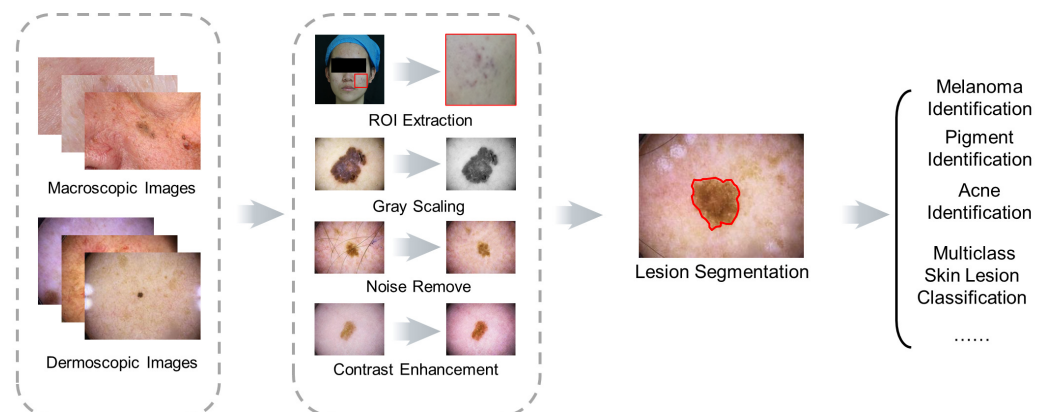
Figure 2. Schematic diagram of skin image diagnosis.

2. Skin Lesion Datasets and Image Preprocessing

2.1. Skin Lesion Datasets

To develop skin diagnosis models, various datasets of different sizes and skin lesion types have been created by educational institutions and medical organizations. These datasets can serve as platforms to educate the general public and as tools to test newly developed diagnosis algorithms.

The first dermoscopic image dataset, PH2, had 200 carefully selected images with segmented lesions, as well as clinical and histological diagnosis records. The Med-Node dataset contains a total of 170 macroscopic images, including 70 melanoma and 100 nevus images, collected by the Department of Dermatology, University Medical Center Groningen. Due to the limited image quantity and classes, these labeled datasets are often used together with others in the development of diagnosis models.

The most popular skin disease detection dataset, ISIC Archive, was created through the collaboration of 30 academic centers and companies worldwide with the aim of improving early melanoma diagnosis and reducing related deaths and unnecessary biopsies. This dataset contained 71,066 images with 24 skin disease classes [3], but only 11 of the 24 classes had over 100 images. The images were organized by diagnostic attributes, such as benign, malignant, and disease classes, and clinical attributes, such as patient and skin lesion information, and image type. Due to the diversity and quantity of images contained, this dataset was used to implement the ISIC challenges from 2016–2020 and dramatically contributed to the development of automatic lesion segmentation, lesion attribute detection, and disease classification [3].

The first ISIC 2016 dataset had 900 images for training and 350 images for testing under 2 classes: melanoma and benign. The dataset gradually increased to cover more disease classes from 2017 to 2020. Images in ISIC 2016–2017 were fully paired with the disease annotation from experts, as well as the ground truth of the skin lesion in the form of binary masks. The unique information included in the ISIC dataset is the diameter

of each skin lesion, which can help to clarify the stage of melanoma. In the ISIC 2018 challenge, the labeled HAM 10,000 dataset served as the training set, which included 10,015 training images with uneven distribution from 7 classes. The BCN 20,000 dataset was utilized as labeled samples in the ISIC 2019 and ISIC 2020 challenge, which consisted of 19,424 dermoscopic images from 9 classes. It is worth noting that not all ISIC datasets were completely labeled in the ISIC 2018–2020 datasets.

To prompt research into diverse skin diseases, images from more disease categories have been (and continue to be) collected. Dermofit has 1300 macroscopic skin lesion images with corresponding segmented masks over 10 classes of diseases, which were captured by a camera under standardized conditions for quality control [14]. DermNet covers 23 classes of skin disease with 21,844 clinical images [11]. For precise diagnosis, melanoma and nevus can be further refined into several subtypes. For example, the EDRA dataset contains only 1011 dermoscopic images from 20 specific categories, including 8 categories of nevus, 6 categories of melanoma, and 6 other skin diseases. The images under melanoma were further divided into melanoma, melanoma (in situ), melanoma (less than 0.76 mm), melanoma (0.76 to 1.5 mm), melanoma (more than 1.5 mm), and melanoma metastasis.

Table 1 summarizes these datasets according to image number, disease categories, number of labeled images, and binary segmentation mask inclusion. It can be seen that the ISIC Archive is the largest public repository with expert annotation from 25 types of diseases, and PH2 and Med-Node are smaller datasets with a focus on distinguishing between nevus and melanoma.

Table 1. Summary of publicly available skin lesion datasets.

Dataset	Image Number	Disease Category	Labeled Images	Segmentation Mask
PH2 [8]	200	2	All	No
Med-Node [12]	170	2	All	No
ISIC Archive [3]	71,066	25	All	No
ISIC 2016 [15]	1279	2	All	Yes
ISIC 2017 [16]	2600	3	All	Yes
ISIC 2018 [17]	11,527	7	10,015	Yes
ISIC 2019 [18]	33,569	8	25,331	No
ISIC 2020 [6]	44,108	9	33,126	No
HAM 10,000 [9]	10,015	7	All	No
BCN 20,000 [10]	19,424	8	All	No
EDRA [19]	1011	10	All	No
DermNet [11]	19,500	23	All	No
Dermofit [14]	1300	10	All	No

Developing a representative dataset for accurate skin diagnosis requires tremendous manpower and time. Hence, it is common to see that one dataset is utilized for multiple diagnosis tasks. For example, the ISIC datasets can be used to identify melanoma or classify diverse skin lesions, while some types of skin lesions have more representative samples than others. Such an uneven data distribution may hinder the ability of developed diagnostic models to handle a desired variety of tasks.

There are no clear guidelines or criteria on the dataset selection for a specific diagnosis task; moreover, there is little discussion on the required size and distribution of a dataset. A small dataset may lead to insufficient model learning, and a large dataset may result in a high computational workload. The trade-off between the size and distribution of the dataset and the accuracy of the diagnosis model should be addressed. On the other hand, it is hard to identify diseases by a diagnosis model for unseen diseases, or diseases with similar symptoms. These intricate factors may restrict the accuracy and, hence, adoption of diagnosis models developed using publicly available datasets in clinical diagnoses.

2.2. Image Preprocessing

Hair, illumination, circles, and black frames can often be found in images in the ISIC datasets, as seen in Figure 2. Their presence lowers image quality and, therefore, hinders subsequent analysis. To alleviate these factors, image pre-processing is utilized, which can involve any combination or all of the following: resizing, grayscale conversion, noise removal, and contrast enhancement. These methods generally improve the diagnosis efficiency and effectiveness, as well as segmentation and classification processes.

Image resizing is a general go-to solution for managing images with varied sizes. With this method, most images can be effectively resized to similar dimensions. Moreover, DL methods can handle images with a broad range of sizes. For example, CNNs have translation invariance, which means that the trained model can accommodate varied image sizes without sacrificing performance [20]. Thus, researchers can confidently tackle skin classification and segmentation tasks under diverse image scales. Resizing images involves cropping irrelevant parts of images and standardizing the pixel count before extracting the region of interest (ROI). This can shorten the computational time and simplify the following diagnosis tasks. The resized images are often converted to grayscale images using grayscale conversion to minimize the influence caused by different skin colors [21]. To eliminate various noise artifacts, as shown in Figure 3, various filters can be introduced [22]. A monomorphic filter can normalize image brightness and increase contrast for ROI determination. Hair removal filters are selected according to hair characteristics. A Gaussian filter, average filter, and median filter can be utilized to remove thin hairs, while an inpainting method can be used for thick hairs and a Dull Razor method for more hair [23,24].

Some ISIC images have colored circles, as shown in Figure 3c, which can be detected and then refilled with healthy skin. The black framing seen in the corners of Figure 3d is often replaced using morphological region filling. Following noise removal, contrast enhancement is achieved for the ROI determination using a bottom-hat filter and contrast-limited adaptive histogram equalization. Face parsing is another commonly used technique to remove the eyes, eyebrows, nose, and mouth as part of the pre-processing process prior to facial skin lesion analysis [25].

(a)Hair   (b)Illumination   (c)Circles   (d)Black frame
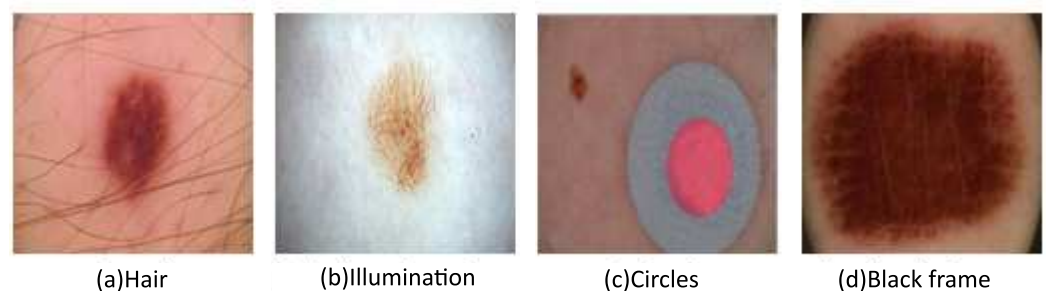
Figure 3. Images with various types of noise from the ISIC archive.

2.3. Segmentation and Classification Evaluation Metrics

The performance of skin lesion segmentation and classification can be evaluated by precision, recall (sensitivity), specificity, accuracy, and F1-score, which are from Equation (1) to Equation (5) [26].

$$Precision = \frac{N_{TP}}{N_{TP} + NFP} \tag{1}$$

$$Recall = \frac{N_{TP}}{N_{TP} + NFN} \tag{2}$$

$$Specificity = \frac{N_{TN}}{N_{TN} + NFP} \tag{3}$$

$$Accuracy = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{TN} + N_{FP} + N_{FN}} \qquad (4)$$

$$F1 = 2 \times \frac{prcision \times recall}{presision + recall} \qquad (5)$$

where N_{TP}, N_{TN}, N_{FP}, and N_{FN} are defined as the number of true-positive samples under each class, true-negative samples under each class, false-positive samples under each class (pixels with wrongly detected class), and false-negative samples (pixels wrongly detected as other classes), respectively.

Other metrics for classification performance measurements include AUC and IoU. AUC stands for the area under the ROC (receiver operating characteristic) curve, and a higher AUC refers to a better distinguishing capability. Intersection over union (IoU) is a good metric for measuring the overlap between two bounding boxes or masks [27]. Table 2 provides a summary of the segmentation and classification tasks, DL methods, and the corresponding evaluation metrics.

Table 2. Summary of tasks, DL methods, and the evaluation metrics. **Jac**: Jaccard index, **Acc**: accuracy, **FS**: F1-score, **SP**: specificity, **SS**: sensitivity.

Task	DL Methods	Metrics						Ref
		Jac	Acc	FS	SP	SS	Dice	
Skin Lesion Segmentation	SkinNet	✓				✓	✓	[28]
Skin Lesion Segmentation	FrCN	✓	✓	✓				[29]
Skin Lesion Segmentation	YOLO and Grabcut Algorithm	✓	✓		✓	✓		[30]
Skin Lesion Segmentation and Classification	Swarm Intelligence (SI)	✓	✓	✓	✓	✓	✓	[31]
Skin Lesion Segmentation	UNet and unsupervised approach	✓					✓	[32]
Skin Lesion Segmentation	CNN and Transformer						✓	[20]
Skin Lesion Classification	VGG and Inception V3		✓	✓				[33]
Skin Lesion Classification	CNN and Transfer Learning		✓	✓	✓	✓		[34]
Melanoma Classification	ResNet and SVM		✓					[35]
Melanoma Classification	Fast RCNN and DenseNet		✓					[36]

3. Skin Lesion Segmentation Methods

Segmenting the ROI is a crucial step in the diagnosis of skin lesions, directly impacting the informative feature extraction process and accurate classification.

3.1. Traditional Segmentation Methods

Traditional image segmentation methods use pixel, region, and edge-based approaches to extract skin lesions from images. Pixel-based segmentation methods, such as binary or Otsu thresholding, can outline each pixel into two categories (i.e., healthy skin or skin lesion). However, these methods can generate discontinuous results, particularly from dermoscopic images, mainly due to low contrast and smooth transitions between lesions and healthy skin. Region-based segmentation can identify and combine adjacent pixels to form skin lesion regions by merging or growing regions. Merging combines adjacent pixels with similar intensity together, while growing starts from a point and checks nearby pixels to expand region coverage [37,38]. The difficulties faced in implementing these methods lie in the variety of colors and textures of individual skin lesions. Edge-based segmentation methods, such as the watershed algorithm [39], utilize intensity changes between adjacent pixels to outline the boundary of a skin lesion. These methods are susceptible to noise, such as hair, skin texture, and air bubbles, which can lead to convergence, especially around noisy points, and produce erroneous segmentation results. In general, traditional

segmentation methods often struggle to achieve accurate results when segmenting images with noise, low contrast, and varied color and texture.

Neural networks (NNs), evolutionary computation, and fuzzy logic have been proposed as methods to segment the ROI based on learning, natural evolution, and human reasoning. These methods can be used individually or in combination to achieve better performance [22]. For example, the fuzzy method has been applied with both splitting and merging techniques to segment dermoscopic images [40]. This combination generated unsupervised perceptual segmentation using fused color and texture features. Later, Devi developed an automatic skin lesion segmentation system using fuzzy c-means (FCM) clustering together with histogram properties [41]. The histogram property was used to select the number of clusters, and the color channel for FCM clustering (hue, saturation, value) was determined based on individual entropy. This system can effectively segment the lesion regions from normal skin automatically with an accuracy of 95.69% when compared with traditional methods. Moreover, there is still much space for advancement in terms of accuracy.

3.2. DL Skin Lesion Segmentation Methods

DL approaches are capable of learning hierarchical features and extracting high-level and powerful characteristics from images. Both supervised and unsupervised DL approaches are used for effective computational segmentation. For example, Vesal proposed a supervised method called SkinNet, a modified version of U-Net, to segment skin lesions in the ISBI 2017 dataset, which is a subset of the ISIC archive that contains a larger number of images from seborrheic keratosis, benign nevus, and melanoma [28]. Al-Masni proposed a full-resolution convolutional network (FrCN) for skin lesion segmentation [29]. This method can directly learn the full-resolution features of each individual pixel of the input data without pre- or post-processing. The segmented skin lesions with fine boundaries were reported when compared with the results from other DL approaches, such as the fully convolutional network, U-Net, and SegNet under the same conditions and with the same datasets.

Multiple DL methods can be combined to search for better solutions. Ünver combined the YOLO and GrabCut algorithms to detect lesion locations using the PH2 and ISBI 2017 datasets [30]. This method was capable of processing dimension-independent images, leading to high-resolution segmentation results. A hybrid DL approach named the Grasshopper Optimization Algorithm (GOA) based on K-means was used to extract the ROI from dermoscopic images following hair removal and contrast enhancement [31]. This hybrid approach achieved better segmentation accuracy when compared with traditional K-means for the ISIC 2017, ISIC 2018, and PH2 datasets.

Due to the tedious process required to obtain pixel-level annotations in skin lesion images, Ali compared the performance of the supervised U-Net and unsupervised methods using F1 scores and IoU [32]. As expected, the supervised method showed much better segmentation accuracy, while the unsupervised method was reported as a good approach when facing a shortage of image annotations.

To balance segmentation performance and annotation workload, a semi-supervised segmentation method was developed. It applied a combined CNN/Transformer model followed by a decoder module to learn a semantic segmentation map and enrich the encoder block by learning an auxiliary task during training [20]. The model's performance demonstrated better results when compared with the unsupervised methods in segmenting images from the ISIC 2017, ISIC 2018, and PH2 datasets.

4. Skin Lesion Classification

Generally, classification tasks in skin lesion analysis can be categorized into binary classification, which aims to identify melanoma for early treatment, and multi-class classification, which is used to identify a variety of skin diseases. Different skin lesions may exhibit similar sizes, textures, colors, and shapes, and there can be significant levels of correlation

between skin lesions, which often poses challenges in classification. Both traditional ML methods and DL methods can be used; however, the former requires feature extraction and selection.

4.1. Feature Extraction and Selection

Feature extraction and selection aim to identify an optimal feature set with excellent discrimination capability in this case to classify skin lesion images. Expert knowledge and clinical experience are preferred, which may effectively and efficiently guide feature extraction and selection. Image processing algorithms can generate features useful for classification, but the effectiveness of extracted features relies on the disease symptoms and diagnosis tasks at hand. To yield accurate melanoma diagnosis, various methods are explored to extract color, texture, edge, and shape features. In the popular ABCD rule, features are selected based on the knowledge that melanoma has an asymmetrical shape, an irregular border, multiple colors and shapes, and a diameter larger than 6 mm. Specifically, A stands for asymmetry, B stands for border, C stands for color, and D stands for diameter [42]. Similar features are extracted using the CASH mnemonic, i.e., color, architecture, symmetry, and homogeneity [43]. In addition, a three-point checklist method was proposed based on the asymmetry of color and structure, atypical network, and blue–white structures observed in melanoma diagnosis [44]. This method was further extended to a seven-point checklist by adding the size of the lesion, itch, or altered sensation of the lesions [45].

The color features can be extracted from the HSV space for hue, saturation, and lightness, the YCbCr space for illumination, and the blue-difference and red-difference chroma and grayscale space. Texture features extracted from grayscale images can reflect the surface roughness of the lesions, such as through the gray-level co-occurrence matrix (GLCM), which includes contrast, correlation, energy, and homogeneity [46], and the neighborhood gray-tone difference matrix, which includes coarseness, busyness, complexity, and texture strength [47]. Additionally, the histogram of oriented gradients (HOG) can be adopted to judge the shape and edge of skin lesions using uniformity or regularity [48].

From the above discussion above, it can be seen that many image features to be extracted are determined from expert knowledge and practical experience, but they may not be all-encompassing in representing skin lesions. The same feature set may not perform well under different tasks, and the effectiveness of such features should be carefully chosen for each specific task [49]. For example, the features extracted using the ABCD rule may yield high accuracy in melanoma detection but not in multiple skin disease classification.

Moreover, the existence of redundant or noisy information can produce irrelevant features that degrade overall diagnostic accuracy. Therefore, it is necessary to select representative feature subsets based on the research tasks and available data and eventually build a classification model with fast training, better performance, and lower complexity.

4.2. DL-Based Feature Extract and Selection Methods

To reduce feature dimensions while preserving the original feature space information, researchers have attempted to explore DL models for feature extraction, such as VGG16, VGG19, and Inception V3 [33]. The extracted features are used as inputs of NNs to classify moles as benign or malignant. The features extracted by the Inception V3 model gave the best accuracy compared to the other two VGG networks for the same classification method. DL model-based feature extraction has also been used in multi-class classification. For example, Kassem used the original fully connected layers with GoogLeNet architecture to extract features and fed them into an SVM for an eight-category classification [34].

DL has also been used in feature selection, which integrates deep feature information to generate the most discriminant feature vector for the classification task. For example, the pre-trained ResNet-50 and ResNet-01 were utilized for feature extraction and then the kurtosis-controlled maximal feature (KcMF) approach was utilized for prominent feature selection [35]). The top 70% of features from the two ResNet models were

fused and fed into an SVM for melanoma diagnosis using three datasets HAM 10,000, ISBI 2017, and ISBI 2016. To recognize skin cancer, Khan combined a DenseNet pre-trained convolutional neural network (CNN) model for deep feature extraction and an iteration-controlled Newton Raphson (IcNR) method for feature selection [36]. For effective feature selection, a framework with the entropy-controlled neighborhood component analysis was proposed to select principle features and extricate redundant features [50]. The effectiveness of this framework was validated on four benchmark datasets, including PH2, ISIC MSK, ISIC UDA, and ISBI-2017. Similarly, Amin fused deep features extracted from the pre-trained AlexNet and VGG16 and then applied principal component analysis (PCA) for optimal feature selection when classifying skin images into benign and malignant [51].

The classification results vary with combinations of feature sets and classifiers. Hegde combined color and texture features in three ways and fed them into four classification models [52], the artificial neural network (ANN), linear discriminant analysis (LDA), SVM, and Naïve Bayes. The extracted features were derived from the ABCD rule, HOG, and GLCM, which were then compared individually and in a hybrid way. The LDA classifier using color features and the SVM classifier using texture features gave the best accuracy in both binary and multi-class classification. The features obtained from the application of the ABCD rule were more informative than the others when applied individually. Among the combined texture and color features, the combination of HOG and GLCM was the best choice when classifying using either LDA or SVM [52].

With very limited training data, more representative features were generated using a deep CNN and feature encoding strategy [53]. The generated features could deal with large variations within melanoma classes, as well as small variations between melanoma and non-melanoma classes in this classification task.

4.3. Traditional ML Models for Skin Disease Classification

The traditional ML methods shown in Figure 4 use extracted and selected features as inputs for classification tasks. After feature selection, traditional ML models such as the SVM, Naïve Bayes (NB) classifier, and K-nearest neighbor (KNN) are widely used to classify skin diseases or identify melanoma as opposed to pigmentation. These traditional ML algorithms with standard structures can generally produce good results in terms of both precision and accuracy [54].
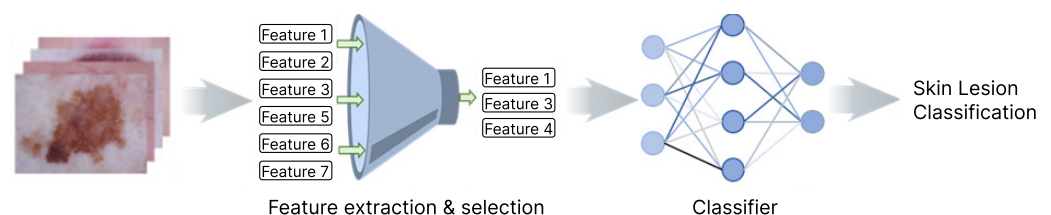
Figure 4. Traditional ML-based skin lesion classification.

The SVM can categorize skin lesions by creating a decision plane to separate different classes based on extracted color, texture, shape, and edge features. The K-nearest neighbor (KNN) differentiates normal skin and skin lesions by comparing the similarity of the input features. Tree-structured classification methods, such as decision tree (DT) and random forest (RF), can handle non-numeric features from the ABCD rule, such as an asymmetrical shape and an irregular border. The NB can classify skin lesion images into their corresponding disease categories based on the highest probability [54].

To compare multi-class categorization performance, Hameed used two strategies to classify skin images [55]. One is a three-category classification: healthy, inflammatory diseases (acne, psoriasis, and eczema), and non-inflammatory diseases (benign and malignant). The other is a six-category classification: healthy, acne, eczema, psoriasis, benign, and malignant. DT, SVM, K-NN, and ensemble classifiers with different kernels were applied to the two classification strategies. The result showed that the classification accu-

racy decreased for all classifiers when the number of categorized classes increased, and quadratic SVM achieved the highest classification accuracy in both strategies. For the same identification task, Hameed also compared the performance of multi-level multi-class classification and single-level multi-class classification [23]. The multi-level multi-class ML classification system achieved better classification accuracy, implemented using both traditional ML and DL.

In summary, the research activities in traditional ML classification models focus on the relevant parameter optimization and input feature selection. Researchers continue to explore the combination of ML algorithms, parameters, and extracted features. Although these activities have been addressed in several studies and successful applications have been made, there is potential to develop new methodologies and improve classification performance.

4.4. Deep Learning Models for Skin Disease Classification

DL has become more sophisticated in its use for image recognition and has achieved increasingly good outcomes [56–59]. As a subcategory of ML, it can automatically discover underlying patterns and identify the most descriptive and salient features for image binary and multi-class classifications, as described by Figure 5.
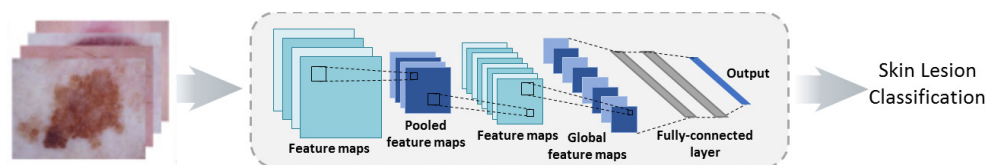
Figure 5. DL-based skin lesion classification.

DL methods are able to process raw image data directly, without the need for a feature preparation step, although they require higher computational costs. Researchers have compared the performance of various DL methods to balance computational cost and classification performance. For example, Ali et al. reported that the DCNN model achieved better performance with less computation time in skin lesion recognition using the HAM 10,000 dataset, compared with the ResNet, AlexNet, MobileNet, VGG-16, and DenseNet models [60]. Additionally, modified DL model architectures have been adopted to enhance multi-class skin lesion classification [34]. One method is to replace the last three layers of the GoogLeNet architecture with a fully-connected SoftMax and a classification output layer for the eight-class classification task. The same task can also be achieved by using the original fully connected layers of the GoogLeNet architecture to extract features and feed them into an SVM. The former achieved better multi-class classification performance for the imbalanced data distribution of the ISIC 2019 dataset.

To take advantage of multiple DL methods, hybrid networks are proposed to diagnose skin diseases. Ahmed combined two CNN architectures, i.e., ResNet and Inception V3 [61], and classified skin lesions into seven types. This ensemble network achieved better accuracy and precision compared to an individual network for the ISIC 2018 dataset. To detect different perspectives of skin lesions, multiple classifiers are used and each classifier focuses on learning one specific aspect. The outputs of these classifiers can be combined to determine the final decision. For example, a two-level system with multiple classifiers was used to exploit the different characteristics of melanoma [62]. The first layer included five individual classifiers: a perceptron combined with color local binary patterns, a perceptron combined with color HOG, one GAN for segmentation coupled with the classic ABCD rule, and two end-to-end individual CNNs (ResNet and AlexNet). A final perceptron was designed to combine these classifiers for melanoma diagnosis. The result exceeded the performance of the individual classifiers when testing with the ISIC 2019 and PH2 databases.

Different loss functions have been explored for the precise detection of melanoma. Since the size of the skin lesion is critical to identifying the disease stage, in two-stage and

three-stage melanoma cancer classifications, Patil combined the CNN architecture with various loss functions, including hinge, loss KL, loss MSE, loss cosine, loss cross-entropy, and similarity measure for text processing (SMTP) [63]. Eventually, an improved CNN architecture with SMTP loss function was chosen due to its better classification performance and efficiency.

The loss function choice can significantly impact the performance of DL models. In image segmentation tasks, commonly used loss functions include (a) Dice loss, which measures the overlap between predicted and ground truth segmentations [64,65], (b) Jaccard loss, which calculates the ratio of intersection over the union of the two sets [66], (c) binary cross-entropy loss, which measures the distance between predicted and true binary masks in binary segmentation [67], and (d) categorical cross-entropy loss for multi-class segmentation under several object classes [65]. In classification tasks, common loss functions include (1) binary cross-entropy loss, which measures the probability distribution over the two classes [68], (2) categorical cross-entropy loss for multi-class classification [68], (3) focal loss, which addresses class imbalance problems by giving more weight to misclassified examples of the minority class, and (4) label smoothing loss, which regularizes the output of the network by adding noise to the ground truth labels. Overall, the loss function should be selected based on the nature of the task and the characteristics of the datasets, which is often an important aspect of designing and training DL models.

As discussed in Section 2.1, some skin lesion datasets are unlabeled, which requires more effective learning rules. A convolutional spiking NN with an unsupervised spike timing–dependent plasticity (STDP) learning rule was developed to distinguish melanoma from melanocytic nevi [69]. As an unsupervised process, this method can learn from small training datasets effectively without the obvious negative influence of overfitting.

In short, research on DL-based skin lesion classification is moving towards exploring available methods, developing new ones, and combining different DL techniques. These efforts include network structure design, layer connection, loss function application, and activation function design. For datasets with large numbers of images per class, DL performs better than traditional ML, and even for datasets with few images, DL can overcome this issue through data augmentation.

There is no solid evidence that DL methods always outperform ML methods in skin disease diagnosis. ML can outperform DL models where representative features are well extracted in the recognition tasks. For example, Alsaade et al. reported that the performance of an ANN model was superior to that of the AlexNet and ResNet50 models when tested on the ISIC 2018 and PH2 datasets [70]. In this study, 216 features were retrieved using the local binary pattern and GLCM, and then fed into the ANN algorithm to detect skin diseases.

5. Current Status, Challenges, and Outlook

Automatic skin lesion image segmentation and classification is a well-established and continuously expanding area of research focus and interest. Its aim is to provide systematic solutions for quantitatively analyzing images, rapidly detecting ROI areas, and reliably diagnosing diseases. With the advancement of image processing methods, these studies offer the potential for developing online diagnosis platforms. However, several challenges and concerns must be addressed before further exploitation.

5.1. Current Research Publication Status

To explain the current research status in skin lesion detection, we conducted a search on the PubMed database on 7 February 2022; this is a free resource that supports the search and retrieval of biomedical and life sciences literature, which can be found in the supplementary material. The research papers were searched using the keywords "skin lesion classification", "skin lesion segmentation", "skin lesion detection", and "skin lesion recognition", while removing papers without "skin" in the title. The search results, shown in Table 3, indicate a dramatic increase in skin diagnosis publications from 2020, with the

number of publications reaching 300 in 2022. Melanoma and skin cancer are identified as the principal areas of focus in skin lesion recognition and have been discussed in 246 papers.

Table 3. Number of publications for skin diagnosis.

Year	2018	2019	2020	2021	2022
Publication No.	179	173	214	258	300

From the collected publications, ML and DL applications in skin disease recognition were discussed in 266 papers as shown in the supplementary document. Out of these, 224 papers mentioned DL application in skin lesion diagnosis, while the remaining 42 research papers used ML methods. The other 28 papers discussed AI applications in skin diagnosis. Thus, the preference of the methodologies adopted by researchers is DL, ML, and other AI methods. The most popular DL method reported is the CNN, which appeared in 43 papers. This coincides with the CNN contribution to image recognition tasks. An interesting observation is the significant number of transfer learning methods found in recent publications, indicating a research trend in skin image diagnosis [71]. As expected, dermoscopic images were the most popular image used in melanoma and skin lesion identification processing.

5.2. Challenges and Outlooks

5.2.1. Macroscopic Images with Robust Diagnosis

Different characteristics of macroscopic and dermoscopic images require particular rules and methods in skin lesion diagnosis, and a direct inference between the two types of images has not been established yet. From the published datasets, one may conclude that dermoscopic images are more commonly collected and, therefore, used more often in computational diagnosis with better segmentation and classification performance. The diagnosis performance of macroscopic or close-up images is not so robust due to the limited number of available images and inconsistent image quality. However, the growing usage of smartphones makes the captured macroscopic images more convenient for online skin lesion diagnosis. To advance recognition model development in skin diagnosis, images collected using smartphones should coincide with the clinician's views during skin lesion examination, and the related features from those images should comprehensively support the diagnosis. Moreover, ethical issues in the dataset collection process should be taken into consideration, especially regarding facial images. Finally, recognition model development should not only target well-known diseases but also progressive or unknown diseases.

5.2.2. Racial and Geographical Biases in Public Datasets

Advances in automatic diagnostics are largely driven by the use of datasets with large repositories of digital images. Public datasets are a compilation of data accessible through an online platform, typically made available by government agencies, academic institutions, or private organizations. Due to the challenges associated with gathering skin lesion images, utilizing publicly available datasets is frequently a more practical alternative to collaborating with medical institutions to obtain such datasets. Public datasets such as the ISIC contain skin disease images mostly from subjects with light-colored skin, collected in the USA, Europe, and Australia. The current recognition models are trained using these unequal distribution datasets. While such recognition models may have outstanding capabilities to recognize lesions on subjects with light-colored skin, their effectiveness is in question when dealing with skin lesion images from subjects from other geographical regions. Moreover, the prevalence and characteristics of skin disease vary with racial and ethnic groups. Traditional ML and DL models fail to offer correct recognition when a test image is from an under-represented skin color group and/or lesion type. More balanced datasets are needed, with clinical data based on gender, age, skin type, and race. This is

critical when using AI diagnosis to improve healthcare in rural areas and increase global access to specialist expertise.

5.2.3. Dataset Characteristics and DL Methods

Along with the rapid development of DL methods, skin lesion diagnosis is moving from an expertise-oriented approach towards a computational intelligence-based approach. When the training datasets are correctly labeled, most DL methods applied for skin lesion analysis are supervised. However, the number of labeled images is not always sufficient for classification model building. Thus, semi-supervised DL methods and weakly-supervised DL methods may be good alternative choices, although they may not achieve comparable performance to supervised methods due to incomplete, inexact, and inaccurate labeled skin lesion datasets. However, in some cases, semi-supervised DL methods can be trained using incompletely labeled datasets, and weakly-supervised methods can handle inaccurate and inexact labeled datasets. Hence, it may be possible to adopt unsupervised DL methods in the cases of unlabeled skin disease images. For example, GAN-based unsupervised methods can learn relatively consistent patterns from a large scale of data without expensive annotation and personal bias [58,72]. Some efforts have been reported on the use of unsupervised methods in biomedical image processing [4].

Meanwhile, unsupervised domain adaptation (UDA) is also a good alternative to reduce the requirement of annotation by transferring the knowledge on labeled data to unlabeled data. Few studies have been reported using fully unlabeled datasets in skin disease recognition, which is likely due to the expected task difficulty and low model performance [71].

Most researchers working on the diagnosis method development are from computer science backgrounds and do not have a fundamental clinical understanding of the diagnosis domain. As such, the current ML and DL models are developed using publicly accessible datasets, which may not be suitable for analyzing real clinical data collected from daily consultations. In other words, the effectiveness of the developed diagnosis models needs to be verified against further independent data not currently available. In addition, future diagnosis frameworks should consider user requirements from the perspectives of both doctors and patients.

6. Conclusions

AI-based skin lesion diagnosis is an increasingly attractive research area, which has been largely driven by the availability of appropriate methods and continually updated abundant datasets. Although relevant topics have been addressed over the last decade, there are still many aspects for investigation and room for improvement.

This paper reviews public skin lesion datasets, the applied image preprocessing methods, and the subsequent skin lesion segmentation and classification methods. The current status, challenges, and outlook in ML-driven skin disease diagnosis are also discussed. Such studies can empower the development of advanced concepts and methodologies. In conclusion, future trends regarding image segmentation and classification of skin lesions require the development of more comprehensive datasets, investigation of more robust models, particularly for macroscopic image recognition, and methods for increasingly reliable automated diagnosis.

Author Contributions: Conceptualization, J.S. and K.Y.; methodology, G.H., C.Z., X.Y.; writing—original draft preparation, G.H., J.S. and M.L.; writing—review and editing: M.L.; supervision, J.S. and K.H.; resources, X.Y. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. ALKolifi-ALEnezi, N.S. A Method Of Skin Disease Detection Using Image Processing And Machine Learning. *Procedia Comput. Sci.* **2019**, *163*, 85–92. [CrossRef]
2. Skin Disorders: Pictures, Causes, Symptoms, and Treatment. Available online: https://www.healthline.com/health/skin-disorders (accessed on 21 February 2023).
3. ISIC Archive. Available online: https://www.isic-archive.com/#!/topWithHeader/wideContentTop/main (accessed on 20 February 2023).
4. Sun, J.; Yao, K.; Huang, K.; Huang, D. Machine learning applications in scaffold based bioprinting. *Mater. Today Proc.* **2022**, *70*, 17–23. [CrossRef]
5. Haenssle, H.A.; Fink, C.; Schneiderbauer, R.; Toberer, F.; Buhl, T.; Blum, A.; Kalloo, A.; Hassen, A.B.H.; Thomas, L.; Enk, 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*, 1836–1842. [CrossRef] [PubMed]
6. Rotemberg, V.; Kurtansky, N.; Betz-Stablein, B.; Caffery, L.; Chousakos, E.; Codella, N.; Combalia, M.; Dusza, S.; Guitera, P.; Gutman, D.; et al. A patient-centric dataset of images and metadata for identifying melanomas using clinical context. *Sci. Data* **2021**, *8*, 34. [CrossRef] [PubMed]
7. Melanoma Skin Cancer Rreport. Melanoma UK. 2020. Available online: https://www.melanomauk.org.uk/2020-melanoma-skin-cancer-report (accessed on 20 February 2023).
8. Mendonça, T.; Ferreira, P.M.; Marques, J.S.; Marcal, A.R.; Rozeira, J. PH 2-A dermoscopic image database for research and benchmarking. In Proceedings of the 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Osaka, Japan, 3–7 July 2013; pp. 5437–5440.
9. 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. [CrossRef] [PubMed]
10. Combalia, M.; Codella, N.C.; Rotemberg, V.; Helba, B.; Vilaplana, V.; Reiter, O.; Carrera, C.; Barreiro, A.; Halpern, A.C.; Puig, S.; et al. Bcn20000: Dermoscopic lesions in the wild. *arXiv* **2019**, arXiv:1908.02288.
11. Dermnet. Kaggle. Available online: https://www.kaggle.com/datasets/shubhamgoel27/dermnet (accessed on 20 February 2023).
12. Giotis, I.; Molders, N.; Land, S.; Biehl, M.; Jonkman, M.F.; Petkov, N. MED-NODE: A computer-assisted melanoma diagnosis system using non-dermoscopic images. *Expert Syst. Appl.* **2015**, *42*, 6578–6585. [CrossRef]
13. Yap, J.; Yolland, W.; Tschandl, P. Multimodal skin lesion classification using deep learning. *Exp. Dermatol.* **2018**, *27*, 1261–1267. [CrossRef]
14. Dermofit Image Library Available from The University of Edinburgh. Available online: https://licensing.edinburgh-innovations.ed.ac.uk/product/dermofit-image-library (accessed on 20 February 2023).
15. Gutman, D.; Codella, N.C.; Celebi, E.; Helba, B.; Marchetti, M.; Mishra, N.; Halpern, A. 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**, arXiv:1605.01397.
16. Codella, N.C.; Gutman, D.; Celebi, M.E.; Helba, B.; Marchetti, M.A.; Dusza, S.W.; Kalloo, A.; Liopyris, K.; Mishra, N.; Kittler, H.; et al. Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging, hosted by the international skin imaging collaboration. In Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2018; pp. 168–172.
17. Codella, N.; Rotemberg, V.; Tschandl, P.; Celebi, M.E.; Dusza, S.; Gutman, D.; Helba, B.; Kalloo, A.; Liopyris, K.; Marchetti, M.; et al. Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic). *arXiv* **2019**, arXiv:1902.03368.
18. ISIC Challenge. Available online: https://challenge.isic-archive.com/landing/2019/ (accessed on 21 February 2023).
19. Kawahara, J.; Daneshvar, S.; Argenziano, G.; Hamarneh, G. Seven-point checklist and skin lesion classification using multitask multimodal neural nets. *IEEE J. Biomed. Health Inform.* **2019**, *23*, 538–546. [CrossRef] [PubMed]
20. Alahmadi, M.D.; Alghamdi, W. Semi-Supervised Skin Lesion Segmentation With Coupling CNN and Transformer Features. *IEEE Access* **2022**, *10*, 122560–122569. [CrossRef]
21. Abhishek, K.; Hamarneh, G.; Drew, M.S. Illumination-based transformations improve skin lesion segmentation in dermoscopic images. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 14–19 June 2020; pp. 728–729.
22. Oliveira, R.B.; Mercedes Filho, E.; Ma, Z.; Papa, J.P.; Pereira, A.S.; Tavares, J.M.R. Computational methods for the image segmentation of pigmented skin lesions: A review. *Comput. Methods Programs Biomed.* **2016**, *131*, 127–141. [CrossRef]
23. Hameed, N.; Shabut, A.M.; Ghosh, M.K.; Hossain, M.A. Multi-class multi-level classification algorithm for skin lesions classification using machine learning techniques. *Expert Syst. Appl.* **2020**, *141*, 112961. [CrossRef]
24. Toossi, M.T.B.; Pourreza, H.R.; Zare, H.; Sigari, M.H.; Layegh, P.; Azimi, A. An effective hair removal algorithm for dermoscopy images. *Ski. Res. Technol.* **2013**, *19*, 230–235. [CrossRef] [PubMed]

25. Zhang, C.; Huang, G.; Yao, K.; Leach, M.; Sun, J.; Huang, K.; Zhou, X.; Yuan, L. A Comparison of Applying Image Processing and Deep Learning in Acne Region Extraction. *J. Image Graph.* **2022**, *10*, 1–6. [CrossRef]

26. Shetty, B.; Fernandes, R.; Rodrigues, A.P.; Chengoden, R.; Bhattacharya, S.; Lakshmanna, K. Skin lesion classification of dermoscopic images using machine learning and convolutional neural network. *Sci. Rep.* **2022**, *12*, 18134. [CrossRef]

27. Gulzar, Y.; Khan, S.A. Skin Lesion Segmentation Based on Vision Transformers and Convolutional Neural Networks—A Comparative Study. *Appl. Sci.* **2022**, *12*, 5990. [CrossRef]

28. Vesal, S.; Ravikumar, N.; Maier, A. SkinNet: A deep learning framework for skin lesion segmentation. In Proceedings of the 2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC), Sydney, Australia, 10–17 November 2018; pp. 1–3.

29. Al-Masni, M.A.; Al-Antari, M.A.; Choi, M.T.; Han, S.M.; Kim, T.S. Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks. *Comput. Methods Programs Biomed.* **2018**, *162*, 221–231. [CrossRef]

30. Ünver, H.M.; Ayan, E. Skin lesion segmentation in dermoscopic images with combination of YOLO and grabcut algorithm. *Diagnostics* **2019**, *9*, 72. [CrossRef]

31. Thapar, P.; Rakhra, M.; Cazzato, G.; Hossain, M.S. A novel hybrid deep learning approach for skin lesion segmentation and classification. *J. Healthc. Eng.* **2022**, *2022*, 1709842. [CrossRef] [PubMed]

32. Ali, A.R.; Li, J.; Trappenberg, T. Supervised versus unsupervised deep learning based methods for skin lesion segmentation in dermoscopy images. In Proceedings of the Advances in Artificial Intelligence: 32nd Canadian Conference on Artificial Intelligence, Canadian AI 2019, Kingston, ON, Canada, 28–31 May 2019; pp. 373–379.

33. Gupta, S.; Panwar, A.; Mishra, K. Skin disease classification using dermoscopy images through deep feature learning models and machine learning classifiers. In Proceedings of the IEEE EUROCON 2021–19th International Conference on Smart Technologies, Lviv, Ukraine, 6–8 July 2021; pp. 170–174.

34. Kassem, M.A.; Hosny, K.M.; Fouad, M.M. Skin lesions classification into eight classes for ISIC 2019 using deep convolutional neural network and transfer learning. *IEEE Access* **2020**, *8*, 114822–114832. [CrossRef]

35. Khan, M.A.; Javed, M.Y.; Sharif, M.; Saba, T.; Rehman, A. Multi-model deep neural network based features extraction and optimal selection approach for skin lesion classification. In Proceedings of the 2019 International Conference on Computer and Information Sciences (ICCIS), Aljouf, Saudi Arabia, 3–4 April 2019; pp. 1–7.

36. Khan, M.A.; Sharif, M.; Akram, T.; Bukhari, S.A.C.; Nayak, R.S. Developed Newton-Raphson based deep features selection framework for skin lesion recognition. *Pattern Recognit. Lett.* **2020**, *129*, 293–303. [CrossRef]

37. Nock, R.; Nielsen, F. Statistical region merging. *IEEE Trans. Pattern Anal. Mach. Intell.* **2004**, *26*, 1452–1458. [CrossRef] [PubMed]

38. Hojjatoleslami, S.; Kittler, J. Region growing: A new approach. *IEEE Trans. Image Process.* **1998**, *7*, 1079–1084. [CrossRef]

39. Levner, I.; Zhang, H. Classification-driven watershed segmentation. *IEEE Trans. Image Process.* **2007**, *16*, 1437–1445. [CrossRef]

40. Maeda, J.; Kawano, A.; Yamauchi, S.; Suzuki, Y.; Marçal, A.; Mendonça, T. Perceptual image segmentation using fuzzy-based hierarchical algorithm and its application to dermoscopy images. In Proceedings of the 2008 IEEE Conference on Soft Computing in Industrial Applications, Muroran, Japan, 25–27 June 2008; pp. 66–71.

41. Devi, S.S.; Laskar, R.H.; Singh, N.H. Fuzzy C-means clustering with histogram based cluster selection for skin lesion segmentation using non-dermoscopic images. *Int. J. Interact. Multimed. Artif. Intell.* **2020**, *6*, 1–6. [CrossRef]

42. Bhati, P.; Singhal, M. Early stage detection and classification of melanoma. In Proceedings of the 2015 Communication, Control and Intelligent Systems (CCIS), Mathura, India, 7–8 November 2015; pp. 181–185.

43. Henning, J.S.; Dusza, S.W.; Wang, S.Q.; Marghoob, A.A.; Rabinovitz, H.S.; Polsky, D.; Kopf, A.W. The CASH (color, architecture, symmetry, and homogeneity) algorithm for dermoscopy. *J. Am. Acad. Dermatol.* **2007**, *56*, 45–52. [CrossRef]

44. Soyer, H.P.; Argenziano, G.; Zalaudek, I.; Corona, R.; Sera, F.; Talamini, R.; Barbato, F.; Baroni, A.; Cicale, L.; Di Stefani, A.; et al. Three-point checklist of dermoscopy. *Dermatology* **2004**, *208*, 27–31. [CrossRef]

45. Argenziano, G.; Catricalà, C.; Ardigo, M.; Buccini, P.; De Simone, P.; Eibenschutz, L.; Ferrari, A.; Mariani, G.; Silipo, V.; Sperduti, I.; et al. Seven-point checklist of dermoscopy revisited. *Br. J. Dermatol.* **2011**, *164*, 785–790. [CrossRef]

46. Oliveira, R.B.; Pereira, A.S.; Tavares, J.M.R. Computational diagnosis of skin lesions from dermoscopic images using combined features. *Neural Comput. Appl.* **2019**, *31*, 6091–6111. [CrossRef]

47. Zakeri, A.; Hokmabadi, A. Improvement in the diagnosis of melanoma and dysplastic lesions by introducing ABCD-PDT features and a hybrid classifier. *Biocybern. Biomed. Eng.* **2018**, *38*, 456–466. [CrossRef]

48. Dalal, N.; Triggs, B. Histograms of oriented gradients for human detection. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–26 June 2005; Volume 1, pp. 886–893.

49. Jie, S.; Hong, G.S.; Rahman, M.; Wong, Y. Feature extraction and selection in tool condition monitoring system. In Proceedings of the AI 2002: Advances in Artificial Intelligence: 15th Australian Joint Conference on Artificial Intelligence, Canberra, Australia, 2–6 December 2002; pp. 487–497.

50. Akram, T.; Lodhi, H.M.J.; Naqvi, S.R.; Naeem, S.; Alhaisoni, M.; Ali, M.; Haider, S.A.; Qadri, N.N. A multilevel features selection framework for skin lesion classification. *Hum. Centric Comput. Inf. Sci.* **2020**, *10*, 1–26. [CrossRef]

51. Amin, J.; Sharif, A.; Gul, N.; Anjum, M.A.; Nisar, M.W.; Azam, F.; Bukhari, S.A.C. Integrated design of deep features fusion for localization and classification of skin cancer. *Pattern Recognit. Lett.* **2020**, *131*, 63–70. [CrossRef]

52. Hegde, P.R.; Shenoy, M.M.; Shekar, B. Comparison of machine learning algorithms for skin disease classification using color and texture features. In Proceedings of the 2018 International Conference on Advances in Computing, Communications and Inform. (ICACCI), Bangalore, India, 19–22 September 2018; pp. 1825–1828.

53. Yu, Z.; Jiang, X.; Zhou, F.; Qin, J.; Ni, D.; Chen, S.; Lei, B.; Wang, T. Melanoma recognition in dermoscopy images via aggregated deep convolutional features. *IEEE Trans. Biomed. Eng.* **2018**, *66*, 1006–1016. [CrossRef]

54. Kassem, M.A.; Hosny, K.M.; Damaševičius, R.; Eltoukhy, M.M. Machine learning and deep learning methods for skin lesion classification and diagnosis: A systematic review. *Diagnostics* **2021**, *11*, 1390. [CrossRef]

55. Hameed, N.; Shabut, A.; Hossain, M.A. A Computer-aided diagnosis system for classifying prominent skin lesions using machine learning. In Proceedings of the 2018 10th Computer Science and Electronic Engineering (CEEC), Essex, UK, 19–21 September 2018; pp. 186–191.

56. Yao, K.; Huang, K.; Sun, J.; Jude, C. Ad-gan: End-to-end unsupervised nuclei segmentation with aligned disentangling training. *arXiv* **2021**, arXiv:2107.11022.

57. Yao, K.; Huang, K.; Sun, J.; Jing, L.; Huang, D.; Jude, C. Scaffold-A549: A benchmark 3D fluorescence image dataset for unsupervised nuclei segmentation. *Cogn. Comput.* **2021**, *13*, 1603–1608. [CrossRef]

58. Yao, K.; Sun, J.; Huang, K.; Jing, L.; Liu, H.; Huang, D.; Jude, C. Analyzing cell-scaffold interaction through unsupervised 3d nuclei segmentation. *Int. J. Bioprinting* **2022**, *8*, 495. [CrossRef]

59. Yao, K.; Huang, K.; Sun, J.; Hussain, A.; Jude, C. PointNu-Net: Simultaneous Multi-tissue Histology Nuclei Segmentation and Classification in the Clinical Wild. *arXiv* **2021**, arXiv:2111.01557.

60. Ali, M.S.; Miah, M.S.; Haque, J.; Rahman, M.M.; Islam, M.K. An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models. *Mach. Learn. Appl.* **2021**, *5*, 100036. [CrossRef]

61. Shahin, A.H.; Kamal, A.; Elattar, M.A. Deep ensemble learning for skin lesion classification from dermoscopic images. In Proceedings of the 2018 9th Cairo International Biomedical Engineering Conference (CIBEC), Cairo, Egypt, 20–22 December 2018; pp. 150–153.

62. Ichim, L.; Popescu, D. Melanoma detection using an objective system based on multiple connected neural networks. *IEEE Access* **2020**, *8*, 179189–179202. [CrossRef]

63. Patil, R.; Bellary, S. Machine learning approach in melanoma cancer stage detection. *J. King Saud-Univ.-Comput. Inf. Sci.* **2022**, *34*, 3285–3293. [CrossRef]

64. Sudre, C.H.; Li, W.; Vercauteren, T.; Ourselin, S.; Jorge Cardoso, M. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In Proceedings of the Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, 14 September 2017; pp. 240–248.

65. Yeung, M.; Sala, E.; Schönlieb, C.B.; Rundo, L. Unified focal loss: Generalising dice and cross entropy-based losses to handle class imbalanced medical image segmentation. *Comput. Med. Imaging Graph.* **2022**, *95*, 102026. [CrossRef] [PubMed]

66. Bertels, J.; Eelbode, T.; Berman, M.; Vandermeulen, D.; Maes, F.; Bisschops, R.; Blaschko, M.B. Optimizing the dice score and jaccard index for medical image segmentation: Theory and practice. In Proceedings of the Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, 13–17 October 2019; pp. 92–100.

67. van Beers, F.; Lindström, A.; Okafor, E.; Wiering, M.A. Deep Neural Networks with Intersection over Union Loss for Binary Image Segmentation. In Proceedings of the ICPRAM, Prague, Czech Republic, 19–21 February 2019; pp. 438–445.

68. Ho, Y.; Wookey, S. The real-world-weight cross-entropy loss function: Modeling the costs of mislabeling. *IEEE Access* **2019**, *8*, 4806–4813. [CrossRef]

69. Zhou, Q.; Shi, Y.; Xu, Z.; Qu, R.; Xu, G. Classifying melanoma skin lesions using convolutional spiking neural networks with unsupervised stdp learning rule. *IEEE Access* **2020**, *8*, 101309–101319. [CrossRef]

70. Alsaade, F.W.; Aldhyani, T.H.; Al-Adhaileh, M.H. Developing a recognition system for diagnosing melanoma skin lesions using artificial intelligence algorithms. *Comput. Math. Methods Med.* **2021**, *2021*, 1–20. [CrossRef]

71. Yao, K.; Su, Z.; Huang, K.; Yang, X.; Sun, J.; Hussain, A.; Coenen, F. A novel 3D unsupervised domain adaptation framework for cross-modality medical image segmentation. *IEEE J. Biomed. Health Inform.* **2022**, *26*, 4976–4986. [CrossRef]

72. Yao, K.; Gao, P.; Yang, X.; Sun, J.; Zhang, R.; Huang, K. Outpainting by queries. In Proceedings of the Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, 23–27 October 2022; pp. 153–169.