Current Medical Imaging, XXXX, *XX*, 000-000

REVIEW ARTICLE

Image Analysis and Diagnosis of Skin Diseases - A Review

Xuhui Li[1,2], Xinyu Zhao[1], Haoran Ma[1] and Bin Xie[1,*]

[1]*School of Automation, Central South University, Changsha, China;* [2]*Hunan Xiangjiang Artificial Intelligence Academy, Changsha, China*

ARTICLE HISTORY

Received: December 20, 2021
Revised: December 09, 2021
Accepted: January 22, 2022

DOI:
10.2174/1573405618666220516114605

Abstract: ***Background:*** Skin disease image analysis has drawn extensive attention from researchers, which can help doctors efficiently diagnose skin disease from medical images. Existing reviews have focused only on the specific task of skin disease diagnosis based on a single medical image type.

Discussion: This paper presents the latest and comprehensive review of image analysis methods in skin diseases, and summarizes over 350 contributions to the field, most of which appeared in the last three years. We first sort out representative publicly available skin datasets and summarize their characteristics. Thereafter, aiming at the typical problems exposed by datasets, we organize the image preprocessing and data enhancement part. Further, we review the single tasks of skin disease image analysis in the literature, such as classification, detection or segmentation, and analyze the improvement direction of their corresponding methods. Additionally, popular multi-task models based on structure and loss function are also investigated.

Conclusions: Challenges involved from the aspects of the dataset and model structure have been discussed.

Keywords: Computer-aided diagnosis, skin disease, deep learning, classification, segmentation, multi-task.

1. INTRODUCTION

Skin diseases are common in people of different regions, races, and age groups in the world, and the incidence is also high. Human skin usually exhibits symptoms including cancer, inflammation and infectious diseases under the induction of adverse environmental factors (such as radiation and lifestyle). There are more than 2,000 types of skin diseases, among which malignant tumors with a high fatality rate are the most harmful, mainly melanoma. Compared with other types of skin diseases, these diseases will easily cause more deaths [1-3]. Early diagnosis is a decisive factor in the recovery of patients with skin diseases. Although the mortality rate of malignant tumors is very high, this type of disease is also one of the easiest cancers to treat if it can be diagnosed at an early stage. Therefore, the research on early diagnosis of skin diseases has aroused widespread concern in the academic and medical circles.

Nowadays, professional dermatologists generally use different non-invasive techniques to obtain medical images corresponding to the diseased area and diagnose the patient's disease. However, the actual clinical application requires a very strict medical diagnosis, among which high medical cost and the subjectivity of the doctor's diagnosis are the two main limitations of medical diagnosis. According to statistics, there are only about 20,000 registered dermatologists in my country, the doctor-patient ratio is extremely wide, and the supply-demand relationship is seriously out of balance. At the same time, the distribution of high-quality medical resources is also very uneven, and there are almost no professional dermatologists in primary medical institutions. Cultivating an excellent doctor requires high costs. Although grassroots doctor training has been carried out all over the country, it is still difficult for doctors or nurses in many grassroots medical institutions to reach the standard level of diagnosis in a short time. Besides, due to the difference in the medical experience of some doctors and the interference of external factors (such as long-term fatigue), personal subjective bias is caused. This kind of doctor's diagnostic subjectivity also greatly affects the correct diagnosis of skin diseases. Considering the limitations of these two aspects, it is necessary to develop some methods to assist doctors in diagnosis.

For some skin cancers, such as melanoma and basal cell tumor, there is predictability due to their regular characteristics. Traditional medical image analysis focuses on the construction of visual features, and then uses the classifier to realize the diagnosis. Generally, the features of lesions involved in the diagnosis of skin diseases cannot be captured by artificial feature extraction algorithms (such as SIFT [4], CN [5]). The feature construction stage of such tasks is usually based on the judgment rules with certain popularity in the industry and practical clinical significance. Some of the rule recognition methods focus on identifying all possible diseases according to the description of category features (such as pattern analysis [6]). Some methods are only used to identify features related to melanoma (such as 7-point

*Address correspondence to this author at the School of Automation, Central South University, Changsha, China; E-mail: xiebin@csu.edu.cn

1573-4056/XX $65.00+.00

checklist [7] and Menzies method [8]). The other is to analyze and diagnose the disease by combining the boundary definition, asymmetry, color status and shape of the lesion (such as ABCD [9] rule and CASH [10] rule). Popular classifiers (such as support vector machines (SVMs) [11], random forest classifier (RF) [12], k-nearest neighbor algorithm (KNN) [13], decision tree (DT) [14] and so on) are generally selected to complete this task. Compared with expert diagnosis, traditional CAD can improve the subjectivity of diagnosis, and the diagnosis result is more objective and reliable. However, using the known evaluation rules will consume a lot of manpower, material resources and time cost in handcrafted data processing, which cannot reach the generalization ability required in the actual situation. Therefore, there is an urgent need for computer-aided diagnosis algorithms with higher accuracy and stronger robustness in the field of medical diagnosis.

With the advancement of artificial intelligence technology, various computer-aided diagnosis (CAD) systems are an effective way to solve the above-mentioned medical diagnosis problems. "Artificial intelligence + Medical", or smart medical for short, has become one of the interdisciplinary subjects with the most development potential and application prospects in the field of artificial intelligence. The practice has proved that CAD has played a very positive role in improving diagnosis accuracy, reducing missed diagnosis, and improving work efficiency.

In recent years, the improvement of GPU computing power has promoted the rapid development of Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) and Auto Encoder (AE), making it possible to train deep neural networks. Deep learning can learn some abstract features that are difficult to be designed artificially, so sometimes it can gain much better performance than artificial features. There are many examples of research using CNN in the field of auxiliary diagnosis of skin diseases. The most representative is the 2017 Stanford team Esteva's research. They used the InceptionV3 network to train a binary classifier for classifying benign and malignant skin cancers, and reached the diagnostic level of professional dermatologists [15]. Encouraged by this, many researches based on CNN have emerged, and CNN has its presence in various medical data fields. The essence of deep learning is driven by data to find better network weights that can represent the characteristics of data.

Most of the data in the field of skin disease diagnosis are presented in the form of images. The visual features of skin lesions are the most critical diagnostic factor in the diagnosis process of skin diseases, so there are currently many studies based on various skin disease images [2, 16, 17]. A dermoscopy image is more suitable for tumor, and it further confirms related clinical diagnosis. It has the characteristics of high permeability and clear background. Most researchers like to use dermoscopy images for research on dermatological auxiliary diagnosis algorithms [18, 19]. The clinical image acquisition method is the first procedure in the actual clinical consultation, so it is suitable for a wide range of diseases. Due to the convenience of its acquisition, clinical images are usually accompanied by great interference, so there are many researches based on such images [20]. In addition, there are some studies based on pathological images, but due to the limitation of the amount of data, such studies are not very sufficient [21].

Image analysis tools have a great influence on skin disease diagnosis. Dermatologists can use these tools to overcome the above problems. These systems usually follow a pipeline: I) image preprocessing, II) lesion boundary detection, III) lesion segmentation and IV) lesion classification. Image preprocessing is a necessary step to process images whose quality is not enough for analysis. As a part of image preprocessing, lesion boundary detection is also the premise of lesion segmentation. Lesion segmentation is the process of obtaining the region of interest, which is necessary for the correct feature extraction and subsequent lesion characterization. Lesion classification is a task in which the diagnosis system automatically extracts features from medical images and gives the results. Therefore, the important part of this paper was to review the literature on several specific tasks of image analysis in the field of dermatology. Fig. (**1**) shows several explorations of deep learning in skin disease image analysis.

In this paper, we did not follow the steps of CAD to carry out a comprehensive review of the latest research on image intelligent analysis technology in the field of skin disease diagnosis. There were other reviews in the field of skin disease diagnosis, and some of the methods proposed were outdated. We reviewed the latest research literature in the field of skin disease diagnosis in the last three years. This was illustrated in Fig. (**2**), so that we could guide readers with more accurate research trends. Different from other reviews that only focused on one type of task or one type of image data acquisition in the field of skin disease diagnosis, the types of research image acquisition we reviewed include dermoscopy images, clinical images, and histopathological images. Similarly, the specific tasks we studied include disease classification, lesion boundary detection, and lesion segmentation. In addition, we also discussed the challenges that still exist in the field of dermatological diagnosis and provided guidelines for dealing with these challenges in the future. Through this paper, people can intuitively understand the development process of image analysis technology in the field of skin disease diagnosis at an appropriate level. Secondly, it can also provide research directions for those who are willing to work further in this field in the future. The remainder of this paper was structured as follows. Section 2 introduced common skin disease image acquisition methods and research-approved public skin disease datasets. In Section 3, we introduced data preprocessing techniques and data enhancement methods. In Sections 4 and 5, we drew on the applicable literature of intelligent image analysis in the skin disease diagnosis task, and introduced the classification, detection, and segmentation tasks in the field of skin disease diagnosis according to the difficulty of entry into the specific task model. Section 6 expands more methods of image intelligent analysis in skin disease diagnosis tasks. Then, in Section 7, we discussed the problems exposed in recent research methods and future research directions. Finally, Section 8 represents our summary of the full text.

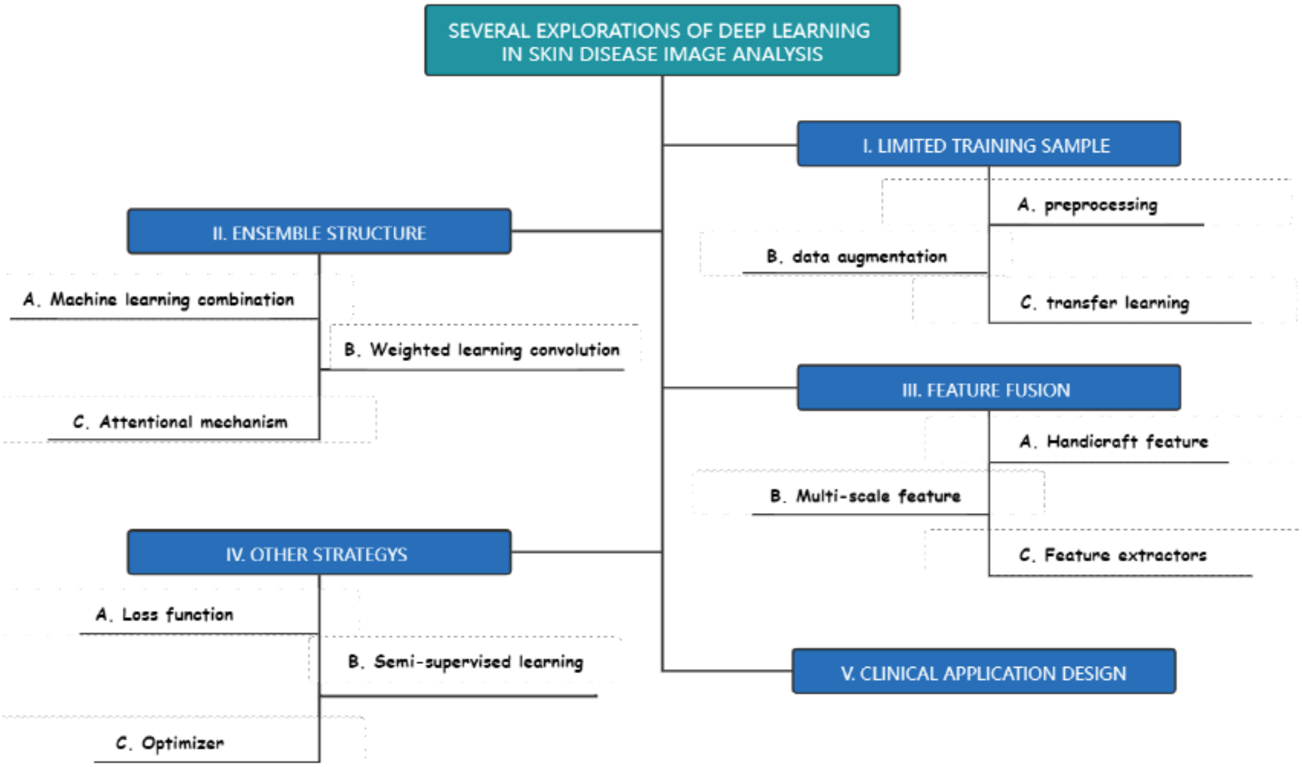

Fig. (1). several explorations of deep learning in skin disease image analysis. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

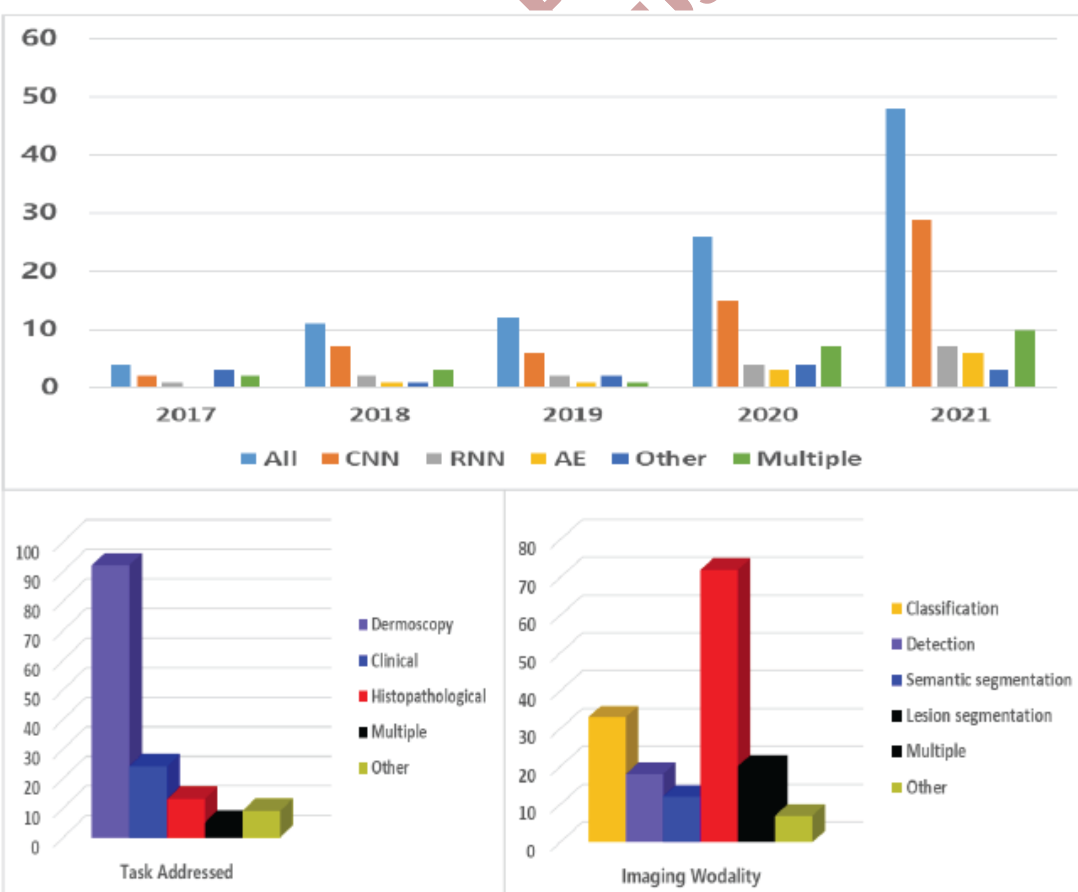

Fig. (2). Breakdown of the papers included in this survey in year of publication, task addressed and imaging modality. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

2. SKIN DISEASE IMAGE ACQUISITION AND DATASET

Most dermatology diagnoses can be performed by visual examination of the skin. Equipment-assisted visual examination is important for dermatologists, because it can provide important help for early accurate diagnosis of dermatology. Different non-invasive imaging techniques are important auxiliary means in the process of examination. Dermoscopy images can enlarge the fine features of skin diseases as needed so that experienced dermatologists can overcome the problems caused by the perception of small-scale skin damage [22]. Clinical images, also called macro images, reproduce the position of clinicians in living organisms that the naked eye sees, and can be used for most skin diseases [23]. Histopathological images can observe the internal structure of the lesions, which is very helpful for the accurate analysis of pathology. The diagnosis information of reflection confocal microscope (RCM) image is like histology, which is based on the morphological and cytological features of the tissue under a microscope. However, compared with histology, the lack of tissue-specific color contrast makes RCM images more difficult to analyze.

2.1. Dermoscopy Image

2.1.1. Description of Dermoscopy Image

In the field of medical diagnosis, most skin diseases can be directly observed and diagnosed by several dermatologists (such as vitiligo, psoriasis and other skin lesions), but it requires years of consultation experience and corresponding medical knowledge. However, the characteristics of some skin diseases (such as cell carcinoma) are difficult to be observed directly, which requires physical means to enlarge the affected skin. Dermoscopy imaging is to optimize the diagnosis technology of skin diseases by using such characteristics.

Dermoscopy imaging technology is one of the most widely used image acquisition methods in dermatology, and it is a non-invasive PSL imaging technique. It allows visualization of its subsurface structure by hand-held incident light magnification equipment (composed of high resolution digital single-lens reflection (DSLR) and optical microscope [24]) and immersion fluid (with a refractive index that makes the skin stratum corneum more transparent to light and eliminates reflection), making the subcutaneous structure easier to see than traditional clinical images [25, 26, 27]. A major change in dermoscopy is the replacement of cross-polarized light with unpolarized light. This allows doctors to capture almost the same image in different situations. However, the "almost" part is the reason for the subtle differences in the visualization of lesions [25, 28], so the application of dermoscopy images will be strictly limited.

Dermoscopy images generally have uniform light and more differentiation, which can not only enlarge the lesion area of the skin but also eliminate some light interference. Fig. (**3**) shows some examples. This technique can assist in diagnosis and analysis by enhancing the characteristics of the skin lesions. It is helpful for professional doctors to identify and detect the morphological structure of skin lesions. However, the two main limitations of dermoscopy are its subjectivity and the requirement of extensive training. It has been demonstrated in the article that dermoscopy may reduce the diagnostic accuracy of inexperienced dermatologists [29].

2.1.2. Common Datasets of Dermoscopy Image

The most used is the public dataset provided by the international skin imaging collaboration (ISIC) [30]. The organization is sponsored by the international society for digital skin imaging, an international organization to improve the diagnosis of melanoma. ISIC archives is a large-scale international repository of dermoscopy images specially set up for clinical knowledge training and image challenge, including tens of thousands of images acquired from international clinical centers. Since the first image challenge was held in 2016, new supplementary categories or new challenges have been added every year. Among them, the image quality is high, and there is no watermark, and the two main types are nevus and melanoma.

The Dermofit dataset [31] is provided by the researchers of Edinburgh University, UK. The data quality is high, and it is widely used by researchers, but it is not free for the public. Dermofit included 10 types of skin diseases: actinic keratosis (45), basal cell carcinoma (239), pigmented nevus (331), seborrheic keratosis (257), squamous cell carcinoma (88), intraepithelial carcinoma (78), pyogenic granuloma (24), hemangioma (97), dermatofibroma (65) and malignant melanoma (76), totaling 1,300.

The published dermoscopy image PH2 dataset [32] is also commonly used as a benchmark dataset for evaluating melanoma diagnostic algorithms, including 160 nevi and 40 melanoma images. Each image is accompanied by comprehensive metadata, including medical segmentation mask, clinical diagnosis (all) and histological diagnosis (part) of the lesion, and medical annotation of several dermoscopy criteria.

The above publicly available datasets for skin diseases are listed in Fig. (**4**). Observing the several dermoscopy image datasets mentioned above, we can draw the following conclusions: 1) Since the acquisition conditions of dermoscopy images are relatively difficult, the total amount of data in the published datasets is not very large. In the context of machine learning algorithms, if it is directly applied to the segmentation or classification of skin lesions, the training and learning of model parameters are likely to be over-fitted, and the diagnostic effect will be greatly affected. 2) There are too few types of skin diseases in the public dataset, resulting in Based on the existing research tasks of the public dataset, the research objects are mostly melanoma-based skin diseases. In the actual clinical situation where there are many types of subdivisions, it is necessary to grasp the specific disease to provide specific disease diagnosis and treatment; the generalization of the research model effect may not be strong enough. 3) The number of categories in the public dataset is not balanced. It is easy to artificially bring the difficulty of the sample to the model learning. When the model cannot clearly determine the disease category, it will tend to classify the object into a large sample category. At the same time, the feature information contained in categories with a small number of samples has great limitations, resulting in a very low recognition rate of this category.

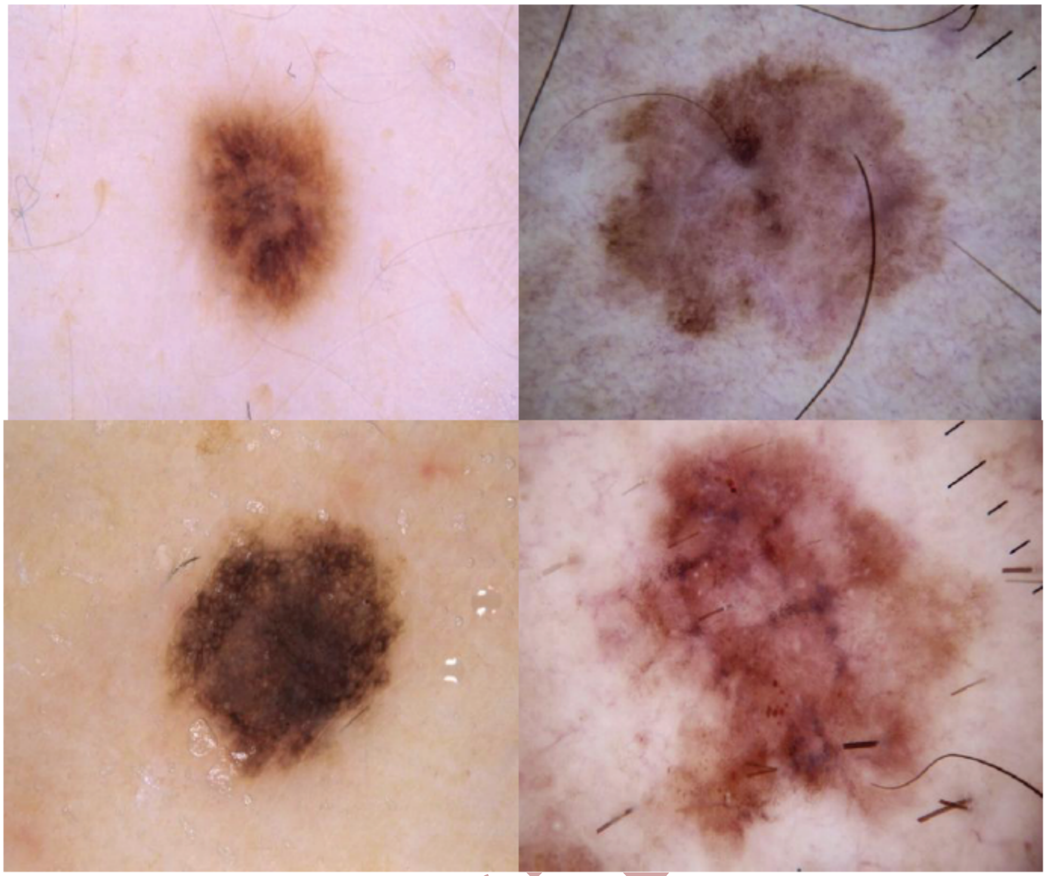

Fig. (3). Dermoscopy images. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

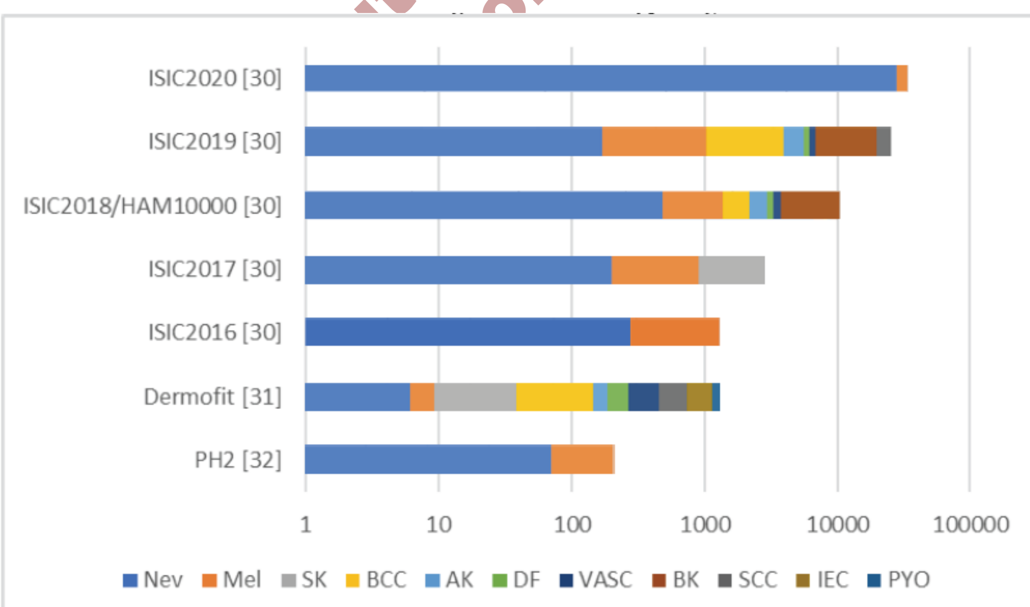

Fig. (4). Common dermoscopy image public datasets, the ordinate is the name of the dataset, the abscissa is the logarithmic scale (lg) to stack the size of the dataset, and the color block length ratio is used to visualize the percentage of various diseases. The diseases include Nevus (Nev), Melanoma (Mel), seborrheic keratosis (SK), basal cell carcinoma (BCC), actinic keratosis (AK), dermatofibroma (DF), hemangioma (VASC), benign keratosis (BK), squamous cell carcinoma (SCC), intraepithelial carcinoma (IEC), pyogenic granuloma (PYO). (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

Fig. (5). Clinical images. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*). (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

2.2. Clinical Image

2.2.1. Description of Clinical Image

Compared with the dermoscopy image, the clinical image acquisition method is much more convenient. The image for clinical diagnosis can be obtained by aiming at the patient's part through a professional camera or smartphone. The purpose of dermoscopy is to enhance the area of the lesion, but the size and location of the lesion will be ignored. Therefore, clinical images can make up for the shortage of dermoscopy, and its convenience will provide an absolute driving force for the application of auxiliary diagnosis in the future, such as remote examination, and patient medical records [33].

However, due to the influence of shooting angle and light intensity, clinical images are almost obtained under different light conditions and uneven focus. This will lead to the external interference of the original clinical image, which is also a problem to be solved in the current research. Fig. (**5**) shows some examples.

2.2.2. Common Datasets of Clinical Image

Dermoscopy images have many large public datasets, and the quality of the diagnostic labels of the datasets can be recognized by the international research team. Comparatively speaking, the quality of clinical image datasets is not as good as that of dermoscopy images (because of the relatively low threshold of acquisition), which leads to a wide range of sources, and many datasets are not so highly recognized in the field of professional research.

MED-NODE dataset [34] includes 70 melanoma and 100 nevus images. Each picture of skin disease is clearly representative and comes from different patients. The dataset has no watermark and is available for free download.

Derm101 dataset [35] is a comprehensive website providing clinicians with various online resources, with the latest treatment suggestions for various skin conditions, as well as high-quality diagnostic images. The dataset contains 22,979 clinical images of 525 categories. Each image contains the basic clinical diagnosis label and the location information of skin lesions. These data are free for academic research. The biggest advantage of this dataset is that there is no watermark information in its image, which is convenient for us to carry out subsequent image analysis and processing. This dataset was not updated starting from December 2019.

SD-198 dataset [36] is a public clinical dermatology image dataset. These images were acquired by digital cameras and mobile phones, covering many patients' conditions, such as age, gender, disease location, skin color and different stages of the disease. The dataset includes 6,584 images from 198 categories. The color, exposure, illumination and scale of the image are different.

Dermnet dataset [37] is a comprehensive website providing a variety of dermatology resources for online medical education through articles, photos and videos. There are nearly 23,000 clinical images (watermarked) and more than 500 kinds of skin diseases in the Dermnet image library. These data are free for academic research. However, each image in this dataset has only a diagnostic tag and no other tags.

Atlasderm dataset [38] is a website of a clinical image database of skin diseases from Brazil, which has 534 categories and 11,009 images. Due to geographical constraints, most of the patients acquired in this dataset are black people, each image has only the label of the diagnosis name, and the image has the watermark of the dataset website.

Danderm dataset [39] is a website for a clinical image dataset of skin diseases from Denmark. There are 91 types of clinical images in this dataset. Most of the patients acquired in this dataset are white people, and only have the label of the diagnosis name, and contain the watermark information of the website. The Atlas holds more than 3,000 clinical pictures and is still expanding.

DermIS dataset [40] is a free public dataset compiled and published by Heidelberg University in Germany. There are 7,172 images in the database, which are divided into 735 categories. In addition to the conventional diagnostic name tag, each image in the datas*et al*so has text description such as race information, lesion location information and age information. The disadvantage of the dataset is that it has too many categories, the number of images in each category is not large, and the image contains the watermark of the dataset website.

Asan dataset [41] was acquired from the Asan medical center facility. After excluding the insufficient postoperative images, 17,125 clinical images were selected from 4,867 patients for 12 kinds of skin diseases. Some of the test set images can be downloaded.

MoleMap [42] is a dataset containing 102,451 images of 25 skin conditions. In particular, the cancerous classification includes melanoma, basal cell carcinoma and squamous cell carcinoma. MoleMap contains a paired combination of clinical images and dermoscopy images. It does not support direct downloading, and data are obtained through cooperation in existing literatures.

In addition, there are some high-quality data sets waiting to be made public, such as XiangyaDerm [20], which has 107,565 images covering 541 skin diseases.

Fig. (**6**) listed several datasets that are relatively frequently used in research. Observing several sets of datasets, we can understand: 1) Most of the datasets are watermarked, which brings additional challenges to preprocessing. 2) The volume of the dataset is much larger than that of the dermoscopy image dataset, of which the MoleMap dataset consists of 102,451 images. Because the source and method of obtaining clinical images are easier, the matching algorithm should be more advantageous in terms of quantity. 2) There are many types of diseases in the dataset. The DermIS dataset has the most 735 types of skin diseases. Ideally, the trained model will have better generalization ability.

2.3. Histopathological Image

2.3.1 Description of Histopathological Image

Histopathological images of skin biopsy are widely used in dermatology. It is very important for dermatologists to understand the potential causes of specific diagnosis through quantitative analysis of digital pathology. Therefore, it is generally believed that biopsy histopathological images are the gold standard for the diagnosis of skin cancer [43]. The preparation and biopsy process of histopathology need to provide a more comprehensive view of the disease tissue with the help of slide, which can retain the potential tissue structure and observe the fine features at the cellular level. Therefore, in the task of automatic image analysis, appropriate use of this micro spatial information can often obtain more specific and more detailed representation of skin diseases from the perspective of pathology. However, the acquisition and processing of histopathological images have the most stringent requirements, and the image itself has a super large order of pixel expression (according to research: the acquisition of patients with 12 to 20 samples will produce 2.5-4 billion pixels [44]), which makes the analysis of these images more difficult and requires higher performance of high-resolution algorithm processing. Fig. (**7**) shown some examples.

All of the public datasets in skin disease task can be found in Table **1**.

2.3.2. Common Datasets of Histopathological Image

Considering the difficulty of application of histopathological images and the complexity of access, large datasets are extremely rare in the research of skin disease auxiliary diagnosis based on histopathological images. In 2020, the international agency for cancer released a dataset of 2,860 images: The Cancer Genome Atlas (TCGA) [45], as shown in Fig. (**8**).

After summarizing the attributes and characteristics of the dataset, the following points were found: 1) The imbalance between the data categories still exists, which is caused by the actual incidence of disease and the actual situation of the medical treatment rate. 2) Compared with other types of images, histopathology images have a high resolution and a large amount of annotation information.

3. IMAGE PREPROCESSING AND DATA AUGMENTATION

In Section 2, we sorted out and compare some publicly available skin datasets of dermoscopy image, clinical image

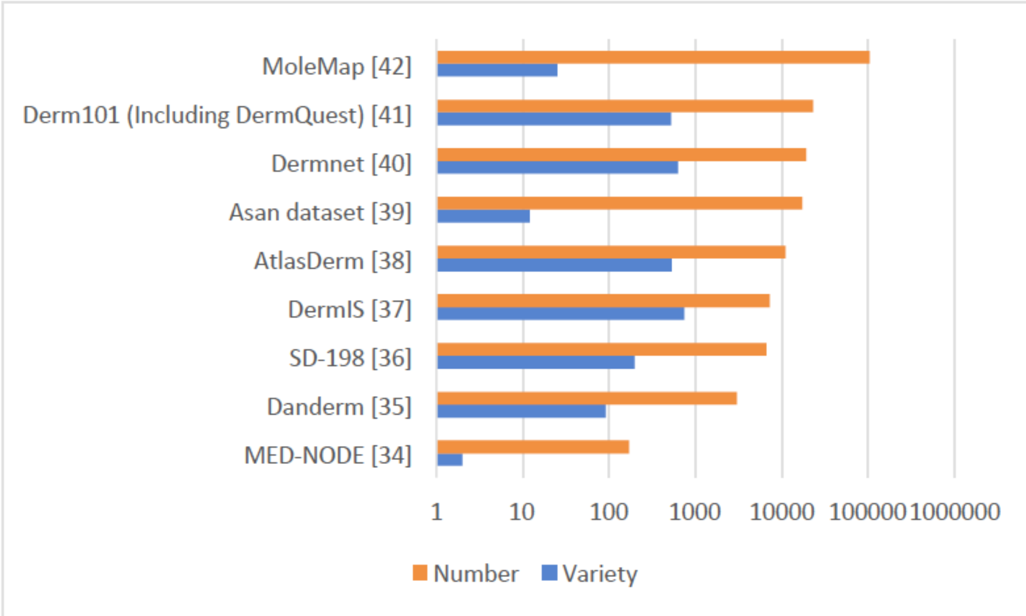

Fig. (6). The published clinical image datasets. The ordinate is the name of the dataset, and the abscissa is a logarithmic scale (lg) to stack the datasets and the number of diseases. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

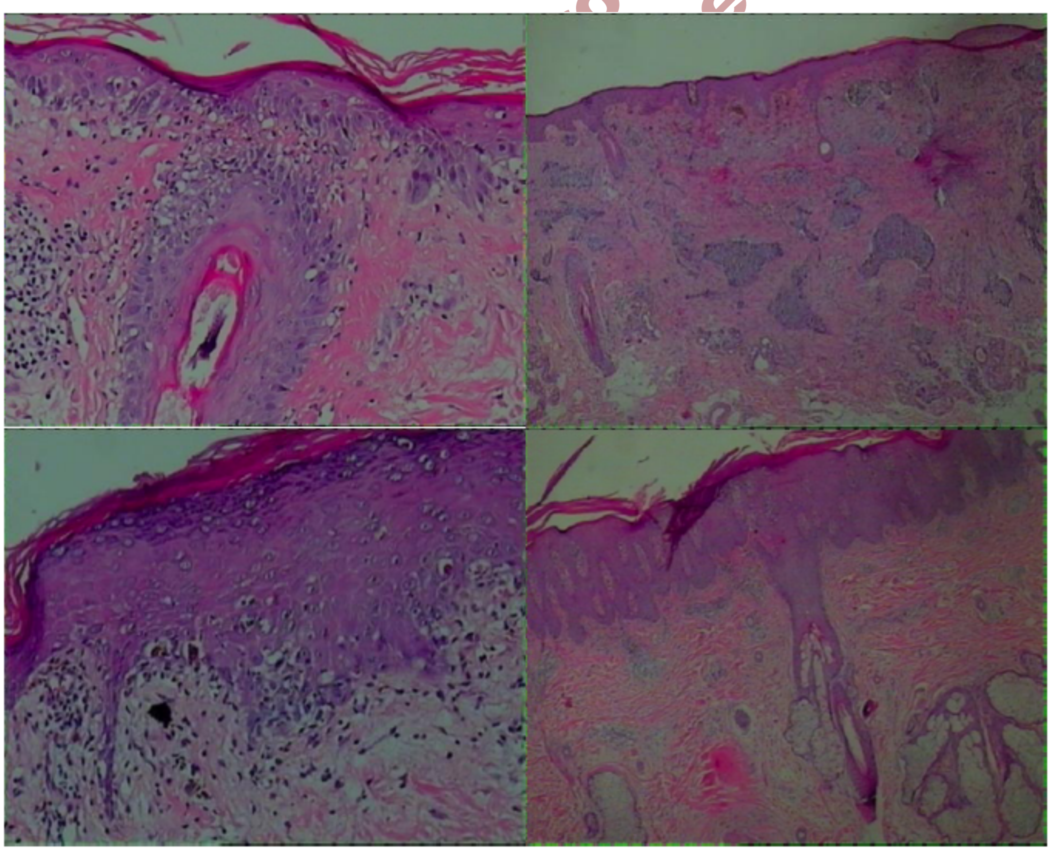

Fig. (7). Histopathological images. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

Table 1. All of the public datasets in skin disease task.

Image	Dataset (Category)	Number	Kind	Typical Disease		Notes
				Nev	Mel	
Dermoscopy image	ISIC2020 [30]	33,126`	2	32,542	584	The image quality is high and there is no watermark, and the two main types are nevus and melanoma.
	ISIC2019 [30]	25,331	8	12,875	4,522	
	ISIC2018/HAM10000 [30]	10,015	7	6,705	1,113	
	ISIC2017 [30]	2,750	3	1,843	521	
	ISIC2016 [30]	1,279	2	1,006	273	
	Dermofit [31]	1,300	10	331	76	It is widely used by researchers including 10 types of skin diseases, but it is not free for public.
	PH2 [32]	200	2	160	40	Each image is accompanied by comprehensive metadata, including medical segmentation mask, clinical diagnosis (all) and histological diagnosis (part) of the lesion, and medical annotation of several dermoscopy criteria
Clinical image	MED-NODE [34]	170	2	\	\	The dataset has no watermark and is available for free download.
	Danderm [35]	3,000	91	\	\	Most of the patients acquired in this dataset are white people, and contain the watermark information of the website.
	SD-198 [36]	6,584	198	\	\	These images were acquired by digital cameras and mobile phones, covering many patients' conditions, such as age, gender, disease location, skin color and different stages of disease
	DermIS [37]	7,172	735	\	\	In addition to the conventional diagnostic name tag, each image in the dataset also has text description such as race information, lesion location information and age information, and the image contains the watermark of the dataset website.
	AtlasDerm [38]	11,009	534	\	\	Most of the patients acquired in this dataset are black people, each image has only the label of diagnosis name, and the image has the watermark of the dataset website.
	Asan dataset [39]	17,125	12	\	\	It was acquired from the Asan medical center facility. Some of the test set images can be downloaded.
	Dermnet [40]	18,974	626	\	\	These data are free for academic research. However, each image in this dataset has only a diagnostic tag and no other tags.
	Derm101 (Including DermQuest) [41]	22,979	525	\	\	Each image contains the basic clinical diagnosis label and the location information of skin lesions. These data are free for academic research. The biggest advantage of this dataset is that there is no watermark information in its image. This dataset was not updated starting from December 2019.
	MoleMap [42]	102,451	25	\	\	It contains a paired combination of clinical images and dermoscopy images. It does not support direct downloading, and data are obtained through cooperation in existing literatures.
Pathological image	The Cancer Genome Atlas	2,875	5	2323	\	Each step in the Genome Characterization Pipeline generated numerous data points, such as: clinical information (e.g., smoking status) molecular analyte metadata (e.g., sample portion weight) molecular characterization data (e.g., gene expression values)

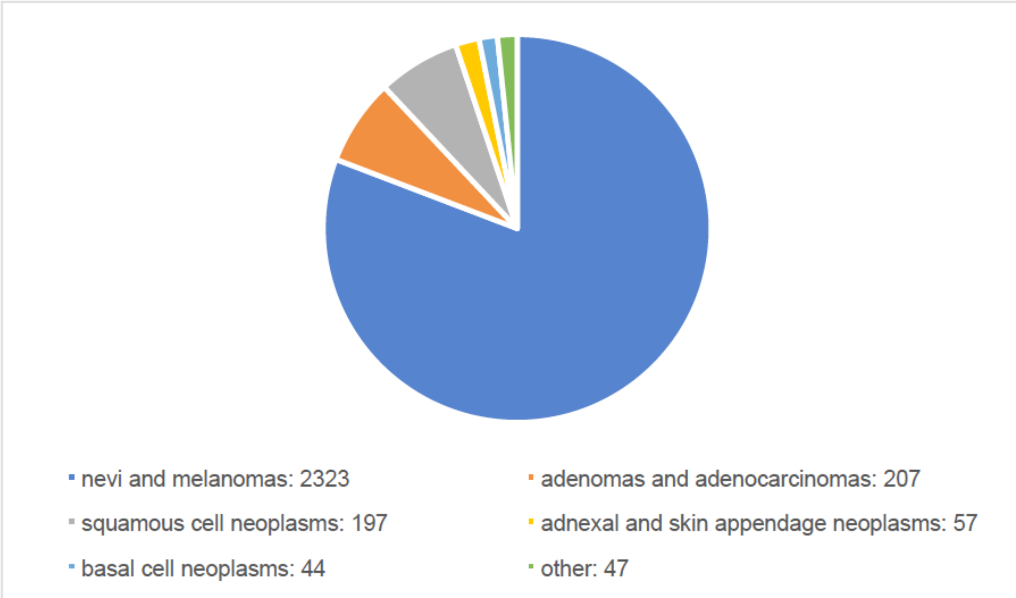

Fig. (8). Histopathology image dataset The Cancer Genome Atlas [45]. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

and histopathological image. In a conclusion, several common problems were exposed there, such as the mixture of chroma and brightness, blurred boundary, artifacts shade. We propose the targeted part of image preprocessing and data enhancement to deal with these problems.

3.1. Image Preprocessing

Because the dermoscopy images and clinical images are all composed of RGB three channels, multi-channel image input is more difficult for the model to process. Therefore, before the dataset image is used as the model input, the necessary channel conversion is always performed. One is to convert a multi-color channels image to a single-color channel. The color information of skin lesions is easier to obtain better contrast difference in a single channel. For example, Poynton *et al*. [46] believed that a separate blue channel could be used for skin lesion detection. There are also some methods to perform appropriate transformation processing on the original image. This practice facilitates further extraction of specific information, such as Pratt *et al*. [47] verified that the combination of Karhunen-LoPve spatial transform variance and maximum variance eliminated the correlation between multi-dimensional data to a certain extent. The channel input processed in this way could provide better diagnostic results.

Although the RGB channel can perform quite well on some tasks, considering the mixture of chroma and brightness, RGB is not a good choice for skin disease image analysis. Different color spaces tend to focus on different visual characteristics. In addition to RGB, commonly used color spaces include hue-based HSV and brightness-based YCbCr. People's normal skin color is greatly affected by brightness in RGB color space. That is to say, after processing in this space, skin color points are discrete points with many non-skin color embedded in the middle, which leads to a difficult challenge to distinguish normal skin parts from abnormal skin parts in RGB color space. If RGB is converted to YCrCb space, Y(brightness) can be ignored. So the space is very little affected by brightness, skin color will produce a good clustering. At the same time, when the channel is reduced to CrCb two dimensions, the sample points will form certain shapes, which is very good for processing pattern recognition. Ahmad *et al*. [48] experimentally proved that better model performance could be achieved by converting data from RGB channel to YCrCb in skin segmentation task. For specific skin disease image analysis tasks, it is sometimes necessary to convert the original image from RGB to YCbCr color space. However, the actual detection ability of the model after the color space conversion is still unstable and easily affected by environmental factors. Many studies have proved that color constancy algorithms, such as grayscale, maximum-RGB, can be used to improve the performance of artificial intelligence algorithms for image classification [49, 50].

It is known that images with blurred boundary of lesions will affect the fit of the model. For example, low-contrast images will increase the difficulty of lesion segmentation and classification. Contrast enhancement is one of the most effective methods in image preprocessing. Recent studies have verified different types of contrast enhancement methods. Gomez *et al*. [51] used histograms to enhance image color channels, which could effectively separate skin lesions and background areas. However, in the actual histogram equalization, a completely flat histogram can rarely be obtained directly. Some studies further discussed adaptive histogram equalization [52, 53]. Since human eyes cannot intuitively capture changes in RGB color space, HSV color space is different, and it will have a much higher match with human eyes. Iyatomi *et al*. [54] proposed a regression model based on HSV color space and verified its excellent effect on color correction. Schaefer *et al*. [55] stated that the combination of gray world algorithm and maximum RGB algorithm can improve the diagnostic effect of the model. Baratu *et al*. [49] analyzed different types of color contrast algorithms, summarized, and showed their comparison results.

In the process of collecting medical image data, artifacts are easily introduced. Artifacts can be roughly classified into two types that are related to the patient (such as hair, skin lines and blood vessels that have a greater impact in clinical images) or machine-related (such as vacuoles, reflections, and black frames generated during dermoscopy image acquisition). These human factors can bring misleading color and texture information and hinder further analysis of the lesion. The detection and removal of hair becomes the most significant step in the handcrafted removal process that has the most significant impact on the model results. One of the difficulties is to remove the cross hairs while introducing as little new interference as possible (such as blur and distortion) to the original image [56]. Bisla *et al*. [57] proposed a data purification method to remove occlusions in the image to achieve a more balanced dataset. Jafari *et al*. [58] investigated the related technical methods of hair removal, and compared the similarities and differences and the scope of applicability. For more accurate segmentation and classification, algorithms including threshold [59, 60], morphology [61, 62, 63] and deep learning [64, 65, 66] should be used to remove hair or other irrelevant things from skin disease images. There were also many studies raising that combining morphology and partial calculus [67, 68] methods for related hair detection tasks had good results.

In order to deal with the many noises and interferences in the dermoscopy image itself, previous studies have tried many different filters (such as Gaussian filtering, mean filtering, median filtering and diffusion filtering) to solve the applicable problems. However, when selecting filters for specific auxiliary diagnosis systems, more consideration is often given to the calculation of the actual model and the time cost.

Clinical images generally have a large size, but the parts involved in the skin lesions are often small or even biased. It is difficult to accurately learn the characteristics of the skin lesions on the original image size during model training. Celebi *et al*. [69] showed that the accurate location of the skin lesions significantly improved the results of the skin lesion segmentation task, and it could help the model to evaluate the size of the skin lesion area. Accurate positioning helps the segmentation method based on contour [70, 71, 72], so that the model can extract more relevant features from the processed image. There were many different methods [73, 74] to solve the problem of skin lesion location. Among them, Celebi *et al*. [74] suggested a threshold-based method to show better results. But it's hard to determine the threshold without a lot of experience. Deep learning can increase the receptive field of the model by using convolution operation, and can strengthen the utilization of macro-lesion location information. Nver *et al*. [75] verified the advantages of deep learning in lesion location. In the research, this part of the task was summarized as the focus boundary detection task, which we introduced in detail in Section 5.

In the public dataset, skin lesion images are acquired with different lighting settings and collection equipment, which may reduce the performance of the artificial intelligence system. Generally, CNNs require inputs with a specific square size (such as 224 × 224 and 512 × 512, etc.). Therefore, it is necessary to resize or crop the images of the research dataset to adapt them to CNNs. It should be noted that resize and crop the image directly to the required size may cause object distortion or loss of a lot of information [76, 77]. The current feasible idea to solve this problem is to resize to a uniform resolution along the shortest side while maintaining the original aspect ratio of the image.

At the same time, in order to accelerate the model convergence and reduce the influence of the variance of data characteristics, the image is normalized by subtracting the average value and then dividing by the standard deviation. The standard deviation is calculated in the entire training subset and then input to CNNs. But under normal circumstances, if the brightness and chromaticity of skin disease images are very different across the entire dataset, subtracting a uniform average value could not be a good way to normalize the illumination of a single image [77, 78]. To solve this problem, the experimental results of Yu *et al*. [77] showed that simply subtracting the uniform average pixel value reduced the performance of the deep network. And it was proposed to realize the normalization of the image channel by calculating the average intensity of the channel of a single image.

3.2. Data Augmentation

In the context of deep learning, the larger the scale and the higher the quality of the data, the better the generalization ability of the model, and the data directly determines the upper limit of model learning. Due to the different incidence of skin diseases and the difficulty in obtaining effective skin lesion images of patients with skin diseases, all existing datasets have some common defects: the total amount of datasets is not enough and the total amount of data between categories is imbalanced. Data enhancement can improve the size and quality of the existing training data set, thereby providing the generalization ability and robustness of the deep learning model, and preventing over-fitting in later applications. In the field of skin disease image diagnosis and analysis, the solution of data enhancement is to reduce many data set limitations of the model through image transformation and image generation technology, such as the ubiquitous data set imbalance phenomenon.

Data augmentation technology can be divided into two categories: supervised and unsupervised. Supervised data augmentation methods mainly include traditional image processing methods based on single sample and multiple samples. The augmentation method of a single sample is to make a certain transformation operation based on the original data. This mainly includes geometric transformation operations (random cut, zoom, rotation, stretch, horizontal and vertical flip, etc.), color transformation operations (color jitter, color space, etc.) and pixel transformation (noise, blur, fusion, etc.) [79]. Nyíri *et al*. [80] concluded that compared to geometric transformations, such as rotation and flipping does not seem to help skin disease image diagnosis at all, but the effect of color transformation was better. Because the geometric transformation has a probability of causing an increase in irrelevant data, these data may have nothing to do with the presentation of the task object, which is completely misleading. Although color transformation is more diversified than geometric transformation, for color-sensitive lesions, rash use of color transformation may cause underfitting of the model [81]. The essence of the augmentation method based on multiple data is to synthesize a new image like the original data. The main methods are SMOTE [82], mixup [83], Sample Pairing [84] and CUTMIX [85] which can continue the discrete original data to fit the new data distribution. Many studies have proved that supervised data augmentation operations improve the diagnosis of skin cancer [86, 87, 88, 89].

The Auto-augmentation [90] using automatic search strategy selection sub- augmentation strategies and Generative Adversarial Network (GAN) [91] are unsupervised data augmentation methods based on machine learning. GAN is a deep learning framework that has aroused interest in the field of medical imaging. Considering that the Auto-augmentation algorithm needs a lot of loss of computing resources and search time, the research team proposed an improved Randaugment [92] and DADA [93]. However, these improved search strategies are still not suitable for the medical field due to actualconditions. In the field of skin disease image analysis and diagnosis, Shen *et al*. [94] proposed a low-cost data augmentation strategy for the implementation of artificial intelligence on mobile devices. The proposed strategy included two consecutive stages: augmentation strategy searched in a discrete optimization search space and random cropped network search of images. It achieved good performance on public datasets at a light-weight search cost. GAN mainly uses the distribution of learning data to randomly generate high-quality pseudo-imaging data consistent with the distribution of the real dataset to overcome the limited dataset [95, 96, 97]. Bissoto *et al*. [98] utilized GAN to generate real synthetic skin cancer lesion images, which solved the problem of lack of annotation data. In the public dataset, the distribution of lesion images will be severely skewed due to the actual prevalence of each category. GAN can be used to generate imaging data for underrepresented skin lesion categories or rare skin cancer categories [99]. In order to alleviate this problem, Bisla *et al*. [100] designed a skin classification system based on GAN to fill rare lesion categories. Combined with the pre-processing algorithm, the system could produce excellent performance that was better than common benchmarks. Compared with the classification task, the demand for labeled data in the segmentation task is more urgent and more difficult to obtain. Gan-generated data can have corresponding pseudo labels at the same time. Pollastri *et al*. [101] presented a new strategy that used GANs to increase the data in the skin lesion segmentation task, which was the basic first step in the automatic detection of melanoma. In the ISIC-2018 dataset [102], skin lesion images are captured at different magnifications or angles or with different cameras. This process is called natural data augmentation. Object detection tasks naturally come with detection anchor frames of different resolutions. It was worth noting that Goyal *et al*. [103] utilized a deep learning architecture called faster region-based CNN (Faster R-CNN) to develop an algorithm to generate augmentation copies of natural data augmentation methods like those used for other skin lesion datasets. For more complex clinical images, Ghorbani *et al*. [104] explored the possibility of using GAN to synthesize clinical images with skin conditions, and the generated images could restore skin conditions with high fidelity. The sources of clinical images are generally not fixed, which can bring additional difficulties to model training. Gu *et al*. [105] used GAN ideas to design a confrontation network, and discussed the advantages of uniform conversion of cross-domain data styles. The confused enhanced image will lead to overfitting of the model in advance and affect the final performance. Yang *et al*. [106] improved the traditional confrontation method and added confrontation loss in the image reconstruction stage to improve the fidelity of the synthesized image. The above operations could be expanded to several times the total amount of data in the original dataset based on the limited amount of data. However, the augmented data essentially has a similar distribution to the original image, which leads to a limited improvement in the generalization performance of the model. The self-attention module can avoid excessively similar features to a certain extent. Abdelhalim *et al*. [107] used GAN combined with the self-attention module to synthesize realistic but completely different skin lesion images from the original images.

4. LESIONS CLASSIFICATION

The diagnosis of skin diseases has always been the first step in the treatment of diseases. But even for experienced dermatologists, the diagnosis process will be very long. At the same time, in the face of diseases with similar visual characteristics of skin lesions, experts are also susceptible to subjective factors to affect the diagnosis. In view of this, a basic task of the CAD system is to realize the recognition and diagnosis of the disease focus, that is, the classification of the disease focus. In this way, it can solve the limitations of high training cost, long diagnosis time, and subjective factors in the diagnosis results encountered by doctors in the diagnosis process. Among them, various machine learning technologies and deep learning models are the mainstream. However, clinical images and dermoscopy images contain many appearance features of the lesion, which are often used as one of the bases for professional doctors to identify the lesion. Since the identification of benign and malignant melanoma has always been a problem for dermatologists around the world, many real medical image datasets are publicly available, so it can often become the verification object for lesion classification tasks in research. However, lesion classification tasks can usually involve more than two lesion categories, namely N classification.

In the related tasks of skin disease diagnosis, the classification task is usually used as the last step of the diagnosis step. The reason why the review is first proposed in this paper is that the network model of the classification task will be more concise and easier to understand, and the detection and segmentation task will be more challenging for researchers. Based on a retrospective review of classification tasks, novices will have more experience in lesion detection and segmentation tasks.

4.1. Classification Algorithm Based on Traditional Machine Learning

Traditional skin disease image classification research is mainly from the lower level of visual features, such as color, texture, shape, etc., to build the descriptors needed by the model, to help automatic and rapid diagnosis in the way of supervised learning. The model obtained by this method has high interpretability and can clearly understand the basis of model judgment. However, due to less parameters than deep learning model, and the ability of handcrafted feature representation is limited. In large-scale data, deep learning model is often not as good as deep learning model, so the number of related researches has become scarce in recent years.

Traditional machine learning classification algorithms are mostly used in dermoscopy images, because the descriptors based on dermoscopy images will be less interfered with by environmental noise, and the related models will perform better. Since skin lesions do not replace normal tissue, but are a way of coexistence, the contrast from RGB will not be obvious. Afza *et al*. [108] utilized the contrast linear stretching process to improve the low-contrast situation at the boundary of the lesion, and proposed entropy to optimize the color feature selection. Their future work was to use deep learning to automate the feature selection process. Mporas *et al*. [109] calculated the statistical values of colors as features, and verified the performance of several classification algorithms. The best effect was the AdaBoost algorithm with an RF classifier.

A single feature sometimes cannot provide all the information of the skin lesions. The ABCD criterion, which contains multiple feature evaluation indicators, is frequently used in the diagnosis of dermatologists and is a set of relatively reliable evaluation rules. Therefore, it is very promising to design a model that can automatically find ABCD features and draw robust conclusions. Monisha *et al*. [110] developed an automatic model to realize lesion segmentation according to ABCD rules, Local binary patterns (LBP) and Gray Level Co-occurrence Matrix (GLCM), and input the result into the Back Propagation (BP) neural network for classification. The ABCD rules contain only the intuitive features that give more information to the model. Chatterjee *et al*. [111] reconstructed a highly robust expert system for benign and malignant identification based on improving the traditional ABCD evaluation indicators. Although clinical images and dermoscopy images have different meaning combinations, they contain the same features of skin damage, such as structure, color, and shape, and can also be used for model research. Based on clinical images, Yang *et al*. [112] studied dermatologists' lesion standards to design several visual features that meet the ABCD criteria, and obtained clinical image classification effects that are not weaker than the CNN model on a small dataset. Considering the different emphasis of visual features, Dhivyaa *et al*. [113] adopted an RF classification algorithm based on a decision-making tree to achieve high performance at a lightweight computational cost.

4.2. Classification Algorithm Based on Deep Learning

Research models based on handcrafted features have a common problem, that is, they rely too much on the current experimental dataset, so the generalization ability of the research model is very insufficient. Therefore, with the rapid development of GPUs, CNN with deep feature learning capabilities has become the first choice for studying model diagnosis effects [15]. The CNN model has the ability of end-to-end learning, that is, the model can directly learn non-representative features, and sometimes it can get far more than artificial features-based visual features. Milton *et al*. [114] compared the disease classification and diagnosis systems of five classic CNNs and obtained the best results on the public dataset. And proposed a larger dataset and better hyperparameters helped to further improve the effect of the model. In actual medical tasks, there are some classification tasks for specific diseases. The performance of different networks varies for different tasks. For example, Singhal *et al*. [115] compared the performance of different networks on the seven classification tasks of skin diseases. In addition to specific diseases, deep learning can also design classification tasks for other specific indicators in the image. Polevaya *et al*. [116] utilized a deep learning network to perform four classification tasks on the main morphological parts of the image end-to-end.

Deep learning methods will bring explosive growth to the demand for data, but data labeling will take up a lot of time and calculation costs. Therefore, relevant literature research in the field is also dedicated to solving this contradiction. Fortunately, data augmentation methods and non-data generation methods (mainly transfer learning [117]) are commonly used to alleviate this problem, which can allow deep learning models to achieve satisfactory results in a limited amount of data. The method of data augmentation has been reviewed in Section 3, such as Qin *et al*. [118] utilized GAN to generate high-quality lesion images, which effectively improved the effect of the classification model.

In addition to data enhancement methods based on the goal of data generation, a large part of the research literature affirms the feasibility of using non-data methods to deal with data problems. Transfer learning is to train a model on a large-scale dataset (such as ImageNet [119]) and transfer its knowledge (weights) to a smaller target dataset. The pre-trained model is a form of transfer learning. The model pre-trained on the ImageNet dataset can bring performance improvements on different modal data in the field of skin disease diagnosis. Because there is some general knowledge in image-based disease recognition tasks and imagenet data set classification. This knowledge can help the model to better converge on new tasks. Joanna *et al*. [120] proposed using a pre-trained VGG-19 network for preoperative melanoma thickness assessment. Experimental results show that the algorithm can achieve the most advanced melanoma thickness prediction results on dermoscopy images. Similarly, Hekler *et al*. [121] utilized a pre-trained ResNet-50 network to diagnose moles and melanomas on histopathological images. A more specialized and more suitable data set would be of great benefit to model training. In cooperation with related institutions, Xie *et al*. [20] built a benchmark dataset for clinical skin diseases and verified the performance of the pre-trained InceptionResNetV2 model, which could achieve good generalization on 107,565 skin disease datasets with a total of 541 categories. Medical images in the field of skin diseases are visually irregular, so the pre-trained models based on ImageNet are not applicable. Compared to directly initializing the weight parameters trained on ImageNet, Gu *et al*. [105] proposed the concept of progressive transfer learning. That was, between ImageNet and the target dataset, by using another skin dataset as an intermediate dataset to perform a first-step fine-tuning, the effect was far better than one-step transfer learning. The deep learning model relies heavily on the data set, and the data set trained on the original data set is difficult to maintain stable performance on the new data set. In the diagnosis of skin diseases, dermatoscopic images are easier to obtain higher performance than clinical images. In order to improve the generalization of CNN at the cross-domain level, Brinker *et al*. [123] trained a model based on the dermoscopy image to achieve dermatologist-level performance on clinical image classification tasks. This model did not need to be trained on the clinical image dataset, and successfully constructed a simple cross-domain model from diagnostic clinical images.

Throughout the existing research literature, classic CNN models are often used as benchmark models. However, some hyperparameters or modules in these models cannot be adapted to all tasks and cannot be universal. In order to improve the handcrafted trial-and-error tuning process, Kwasigroch *et al*. [122] presented a neural architecture search method to automate this process, and made a lot of efforts to reduce the number of parameters. After the model training is completed, some network structures are assigned very low weights, so deleting these structures can reduce the number of parameters without excessively affecting the performance. Based on the existing CNN models, Muckatira *et al*. [124] improved the performance of the original network by more than 10% by trimming some of the network parameters. Pruning was achieved by resetting the weights below the set threshold to zero in the range of the entire model. As CNN models were further researched and demonstrated, a series of improved CNN models were constructed. The latest advances in deep learning include hole convolution, which was known to have higher accuracy under the same computational complexity than traditional CNN. Ratul *et al*. [125] compared the transfer of four classic network architectures based on hole convolution. Learning effect, the new model had better average accuracy than any known method. However, because CNN directly outputs the lesion class, such a diagnosis generally lacks interpretability. As it happens, image retrieval based on interpretable content is expected to be a supplement to clinical decision-making [126]. Allegretti *et al*. [127] proposed a skin disease image retrieval method. After the embedding layer was connected to the feature layer, the distance loss was used to evaluate the similarity with the target image, which could reach the average level of prediction of a human-computer battle.

The attention mechanism is essential to learn a weight for the input feature vector in order to highlight salient features and suppress irrelevant features. It has been widely used in skin disease classification recently [128]. Barata *et al*. [129] proposed a deep attention model that combined CNN and Long Short-Term Memory (LSTM) with an attention module to help identify relevant areas in skin lesions and guide classification decisions. Aimed at the similarity of images between classes and the emphasis on the key features of skin lesions, Aggarwal *et al*. [130] presented an attention-guided deep CNN to achieve the two classifications of benign and malignant melanoma. The results showed that a careful model could increase the accuracy of a normal CNN model by more than 10%. Similarly, Zhang *et al*. [131] also proposed a novel attention module to use the features learned in the upper layer to generate the attention feature map of the lower layer.

Although the accuracy of research literature on authoritative datasets continues to hit new highs, the actual task requirements will lead to both the early classic VGG and the latest ResNeSt [132] in the diagnosis of specific skin diseases. Different deep neural networks have their own biases on the learned features. One solution to this problem is to train multiple models as feature extractors to learn and evaluate different features of skin lesions, and combine the submitted predictions to generate the final diagnosis. This method uses the ideas of ensemble learning [133] to reduce prediction errors. This integrated method usually produces better results than any single model, and has been applied to the classification of skin diseases. Wang *et al*. [134] tried to use the output vectors of five classic CNN models to achieve seven classifications of skin lesions. In the experiment, the classifier with the best parameters obtained by the Bayesian search algorithm was also used to achieve High efficiency and accuracy. The method proposed by Mahbod *et al*. [135] was composed of multiple groups of CNNs with different architectures focusing on different features, and the advanced features of multiple neural networks were merged in the later stage. Experiments showed that the fused features had better and more adequate feature description capabilities. Considering more model structures, Perez *et al*. [136] repeatedly evaluated the classification of melanoma on public dataset and found 135 models, and found that no matter how they were combined, the integration of multiple models had an advantage over a single model.

Due to the superiority of the integrated method in improving the performance of the model, people have made some new attempts to its combination in specific tasks. Some researches try to divide the skin lesion classification problem into multiple sub-problems, and divide these subproblems into multiple steps to solve them, rather than in one step to improve the classification performance. Skin conditions can first be divided into benign and malignant categories, which has a strong clinical interpretation. In the task of skin seven classifications, Harangi *et al*. [137] introduced the confidence value of the upper level of benign and malignant classification as the relevant normalization coefficient. Two classification tasks with higher quality datasets can often obtain more accurate confidence values, which can improve the category imbalance in the seven classifications. Thanks to the benefits of multi-stage tasks, more manually designed stage tasks are being explored. Barata *et al*. [128] set up a three-stage classification task to help refine the effect of the classification task. Similarly, Hameed *et al*. [138] utilized traditional methods and deep learning methods to verify the excellent capabilities of multi-class and multi-level classification algorithms on datasets from different sources. Furthermore, Mahbod *et al*. [139] fused the feature layer of the model from three stages (cross-validation, image crop size, and model structure), and had excellent classification performance on the public dataset.

This integrated method is basically a later fusion of different features extracted from different network models. Its success has inspired people to perform more hierarchical feature fusion. Tang *et al*. [140] proposed to combine the global and regional features of the image, and integrated four different scales of image input to further improve the classification ability. Different from the integrated learning technology of voting combination, stacking is another technology used to generate the metadata feature space required by the task. Stacking is a framework for integrating several feature extraction models using a single model. Guided by the idea of multi-level classification, Ghalejoogh *et al*. [141] replaced voting modules with stacked modules, and the performance of the model far exceeded the prediction performance of a single classifier. Besides, Zhang *et al*. [19] presented a collaborative deep learning model in which two independent deep CNN models were connected in series to a collaborative network to solve the problem of significant intra-class differences and inter-class similarity. There were many prospective studies that had been proposed. Models of different structures can be stacked together to bring different dimensions of information to the classification model. Among them, Walker *et al*. [142] claimed in experiments that the output after ultrasonication of the image feature layer had high sensitivity to lesions. They suggested that the model of ultrasound output combined with the classifier could evolve into a useful decision support system for all doctors to use, and it also brought a lot of inspiration for subsequent research. The ability of machine learning technology to transform input data into high-level representations has received widespread attention in recent years. Sabbaghi *et al*. [143] utilized the bag-of-feature (BOF) model to cluster the combination of SIFT features and color features, and input them to the stacked sparse autoencoder to complete the skin lesion classification task through an unsupervised scheme.

Also, strategy-level methods also provide a great help in improving model performance. Most of the existing CNN models used in skin disease research use classic loss functions, such as cross-entropy loss functions. For specific skin disease tasks, this may limit the model's ability to further learn and recognize features from skin disease images. Different dimensions of supervision may be more beneficial to model performance in some tasks. In order to solve the above problems, Ahmad *et al*. [144] presented a new framework to try to improve the classification of skin diseases based on the triplet loss function and fine-tuning models. In addition, in response to the dataset problems exposed during the training process, Lin *et al*. [145] described a new loss function to improve the interference of imbalanced datasets and difficult-to-separate samples. In the testing and training stages of the CNN model, different activation functions have different abilities to solve the non-linear factors of the dataset. Therefore, considering the specificity of the model in the dataset, Goceri *et al*. [146] analyzed and compared the clinical image automatic diagnosis effect of the two activation functions on the four network models.

In addition to improving the training level strategy of the loss function, it can also contribute at the preprocessing level. In order to reduce the cost of cost-effective skin data labeling, Shi *et al*. [147] only used active learning of sample selection strategies and further added sample expansion strategies to achieve high-level skin lesion analysis. The experiment used half of the dataset to achieve the most advanced performance on two different tasks. Semi-supervised learning model can make full use of limited marker data and a large amount of original data, which is of great help and improvement to medical image analysis tasks. At the same time, Bdair *et al*. [148] proposed a new semi-supervised learning model, which was a process of generating pseudo-labels of unlabeled samples by sharing the knowledge of training samples, which was more than 15% higher than the baseline. The excellent performance in the existing literature had not been widely used, which may be due to the uncertainty of real data, which could easily lead to a lack of confidence in automatic diagnosis or errors in result interpretation. Designing different tasks on the same data set can further utilize the information of the data. Different from the previous method, Bagchi *et al*. [149] made full use of limited data to improve accuracy. After the diagnosis model, a re-identification network was connected in series to confirm whether the diagnosis model predicted correctly, and other unknown categories of data were used to train the network. Combalia *et al*. [150] utilized Monte Carlo Omission and several estimation techniques that could increase the uncertainty of training data to improve the true diagnostic performance of the classifier.

At present, many studies have obtained results equivalent to the diagnosis level of doctors based on dermoscopy images, but when it comes to clinical images that are easier to obtain, it will inevitably bring more severe challenges. The research results need to be diagnosed as accurately as possible on clinical images before they are implemented in the user experience environment. Compared with dermoscopy images, most studies based on clinical images are collected by the research team itself, so the image quality is different and the difficulty is relatively high. Jinnai *et al*. [151] trained a Faster R-CNN model on a dataset of 3,551 patient clinical images, and obtained a diagnostic effect that was not lost to dermatologists. One of the important significances of clinical image research is its application in hardware platform to achieve the popularization of auxiliary diagnosis. Khamparia *et al*. [152] provided a novel, IoT-driven deep learning framework for melanoma classification in skin disease images, which could be used remotely to assist medical experts in the diagnosis and treatment of skin cancer. Integrating algorithms into mobile devices could also bring artificial intelligence technology to more people who need it. Hameed *et al*. [153] proposed a movable skin damage classification expert system application. The system could quickly realize the four classifications of skin types in clinical images, which not only facilitated patient self-examination, but also simplified the preliminary diagnosis steps of professional physicians to a greater extent, so that experts could focus more on a detailed diagnosis of difficult diseases. However, when the model was tested on multiple smartphone images taken by different cameras under different lighting conditions and distances, it was difficult to achieve the same diagnostic accuracy. For example, the self-acquired skin disease images provided by patients were usually of low quality and were not suitable for auxiliary diagnosis [154, 155].

After people further studied the performance improvement factors of the diagnostic network model, a conclusion was reached: the dataset determines the upper and lower limits of the performance of the diagnostic model to the greatest extent. As mentioned above, researchers have paid attention to the size and source of experimental datasets. Here, the characteristics and fusion of multiple data types will be described. Dermoscopy images can show the standardized field of view of blood vessels, dots and balls on the lower surface of the lesion to provide texture features. Clinical images describe the geometry and color of lesions, which are not so standardized in comparison. It supports more relaxed shooting in different fields of view, but always contains some image artifacts. In addition to image data, the patient's metadata includes other types of information, such as the patient's gender and lesion location. Together, these methods can provide a standard form to describe skin damage, which is the best aid in the actual diagnosis process. Feature fusion methods (such as early fusion and later fusion) have also been widely discussed. Among the few published skin disease diagnosis papers based on multi-modality, these methods are based on post-fusion methods to integrate image modal features extracted by multiple independent networks. In the future, it is expected to be extended to more anatomically significant prior knowledge to help segmentation of medical images. Ge Z *et al*. [156] adopted deep CNN as the benchmark model to build a new Triple-Net, which consisted of two single-mode models and a cross model. The advantages of multi-modal data input were verified in three different feature fusion methods. On the cross-modal dataset, the effect was far better than that of the single-modal model. The authors concluded that combining Class activation map and Bilinear Pooling methods could capture complementary feature information from cross-modal data. But there could be more adaptations in the hyperparameter selection of Triple-Net's loss fusion. Clinical diagnosis of skin diseases will generate a lot of text data, which will be helpful for doctors to make diagnostic decisions. Kawahara *et al*. [157] proposed an integrated neural network designed for multimodal datasets of dermoscopy images, clinical images, and metadata. It classified skin damage diagnosis according to the criteria of all seven-point checklists in a multi-task model. At the end of the article, the visualized corresponding area of the modal was realized to enhance the interpretability. The stage of model fusion is different in different tasks or different types of data, and the design of this part in general studies depends on experience. Nunnari *et al*. [158] explored the effects of two different multimodal fusion methods. The author confirmed that shallow network fusion could avoid the masking of metadata input and achieve better classification results. The CNN architecture proposed by Yap *et al*. [159] accepted simultaneous input of dermoscopy images, clinical images, and patient metadata (including patient age, gender, and lesion location). They proved through experiments that the multimodal method was better than the single-modal method in the five types of skin tumor classification and the binary classification of melanoma. However, it was discussed in the experiment that the patient's metadata did not seem to be able to achieve the expected improvement effect. Future research would explore clinical information with more direct diagnostic significance. Similarly, in a recent study, Pacheco *et al*. [160] combined dermoscopy images and patient clinical information in a deep learning model, achieving an improvement of about 7% in average prediction accuracy. At the same time, it is concluded that clinical information was not helpful for the classification of all lesion types. In the future, this work will focus on more image modalities. Different from the later fusion method, some people believe that the multi-modal image features from each level of the CNN structure will help the classification of multiple types of skin lesions. For example, high-level features focus on the classification of lesions, while low-level features have an advantage in predicting the area of skin lesions. Bi *et al*. [161] introduced a deep hyperconnected CNN with multi-scale attention blocks to integrate the visual characteristics of clinical and dermoscopy lesion images.

In order to facilitate reference to the current literature on the classification of skin lesions, we list Table **2** below by year.

Table 2. Reference for classification method of skin lesions.

Refs.	Year	Dataset (Category)	Model & Method	Remarks
[196]	2016	814 images (2)	BOF + Encoders	Rriplet loss function and fine-tuning models
[198]	2017	MoleMap (15)	VGG-16	Dermoscopy, clinical images, and metadata; integrated neural network
[112]	2018	SD-198	SVM	Designing several visual features which meet the ABCD criteria
[158]	2018	ISIC 2016	ResNet-50	Ultrasonication of the image feature layer
[160]	2018	1,011 cases (5)	Inception V3	Multimodal fusion methods
[166]	2018	2,917 cases (5)	ResNet-50	Combining dermoscopy images and patient clinical information
[110]	2019	\	ANN + BP	Multiple feature evaluation indicators
[114]	2019	ISIC 2018	PNASNet-5-Large	Comparing the five classic CNNs
[116]	2019	\ (4)	VGG-16	End-to-end network on the main morphological parts of the image
[123]	2019	HAM10,000	ResNet50	Basing on the public dermoscopy image

Table (2) contd….

Refs.	Year	Dataset (Category)	Model & Method	Remarks
[120]	2019	\ (3)	VGG-19	A pre-trained VGG-19 network
[121]	2019	595 images (2)	ResNet-50	Pre-trained ResNet-50 network
[20]	2019	107,565 images (541)	InceptionResNetV2	Pre-trained InceptionResNetV2 model
[105]	2019	MoleMap + HAM10,000	Cycle-GAN	Progressive transfer learning
[129]	2019	ISIC 2017, 2018	LSTM+DenseNet-161+ResNet Inception	Deep attention model which combined CNN and LSTM with attention module
[130]	2019	HAM10,000	InceptionV3 + Attention	An attention-guided deep CNN
[131]	2019	ISIC 2017	ResNet-50+ Attention	Using the features learned in the upper layer
[135]	2019	ISIC 2017	4 CNNs	Advanced features of multiple neural networks merged in the later stage
[136]	2019	ISIC 2017	9 CNNs	The integration of multiple models
[147]	2019	3,954 images (2)	Inception V2	Utilizing the BOF model to cluster the combination of SIFT features
[148]	2019	100 images (5)	ResNet	Activating learning of sample selection strategies
[149]	2019	ISIC 2017	ResNet-101	Semi-supervised learning by sharing the knowledge of training samples
[154]	2019	1,856 images (4)	SqueezeNet	IoT-driven deep learning framework
[108]	2020	ISIC 2017	SVM	Contrasting linear stretching process and optimize the color feature selection
[109]	2020	HAM10,000	AdaBoost	Calculated the statistical values of colors as features
[113]	2020	ISIC 2017 + HAM10,000	DT + RF	A RF classification algorithm based on DT
[115]	2020	HAM10,000	Inception ResNet v2	Transfering learning models
[118]	2020	ISIC 2018	ResNet-50 + GAN	Utilizing GAN to generate high-quality lesion images
[122]	2020	ISIC 2019	VGG	Neural architecture search method to automate trial-and-error tuning
[124]	2020	ISIC 2019	ResNet-18	Trimming some of the CNN parameters
[125]	2020	HAM10,000	InceptionV3	Transfer of network architectures based on hole convolution
[127]	2020	ISIC 2019 (retrieval)	ResNet-50	Skin Lesion Diagnosis
[134]	2020	ISIC 2018	SE-Resnext-101	Machine Learning
[137]	2020	ISIC 2018	Inception-v3	Multi-class classification considering a binary classification support
[139]	2020	3,672 images (4)	AlexNet	Using the feature layer of the model from cross-validation, image crop size, and model structure
[140]	2020	ISIC 2018	3 CNNs	Combining the global and regional features
[141]	2020	ISIC 2017	Xception	Replacing voting modules with stacked modules
[145]	2020	PH2, Ganster	4 Classifiers	Two independent deep CNN models were connected in series
[150]	2020	6,144 images (5)	ResNet-152 + Inception-ResNet-V2	Focal loss for dense object detection
[151]	2020	ISIC 2019	3 CNNs	Monte Carlo Omission and several estimation techniques
[152]	2020	ISIC 2019	EfficientNet-B0	Faster R-CNN
[159]	2020	4,732 images (2)	Faster R-CNN	Movable skin damage classification expert system
[161]	2020	ISIC 2019	ResNet-50	CNN
[162]	2020	1,612 images (6)	6 CNNs	Deep hyperconnected CNN with multi-scale attention blocks
[163]	2020	1,011 cases (5)	ResNet	The accurate segmentation of the image
[111]	2021	ISIC 2016, 2017, 2018, PH2	ABCD rules	Expert system for benign and malignant identification
[138]	2021	ISIC 2018	DenseNet-161	Multi-class and multi-level classification algorithms
[150]	2021	38,000 images (>10)	EffecientNet	Meta-ensemble technique

5. LESION SEGMENTATION

Skin damage in dermoscopy images and clinical images is a single bounded area, which is usually distinguished from normal surrounding skin due to different colors or textures. This area is an area of interest that needs further processing in skin disease diagnosis tasks [162]. Lesion segmentation is to separate the focus area of research interest from the background, which helps clinicians perceive the boundary of the lesion. At the same time, the accurate segmentation of the image can provide a region of interest for the later clinical feature segmentation, so the success of image analysis depends on the reliability of the segmentation [163]. Although lesion segmentation is sometimes used to improve several downstream tasks, such as lesion classification, it itself is largely understood as an important and challenging task.

Handcrafted boundary detection mainly considers the obvious visual difference between the lesion area and healthy skin, but more than one type of damage is more often similar. Therefore, a higher level of knowledge of lesion characteristics should be considered [52]. However, the handcrafted boundary detection method not only has a low misdiagnosis rate, but also the doctor's operation is not easy to replicate, and the scope of application is bound to be limited. The segmentation algorithm design of skin disease images has quickly attracted the attention of scholars. Lesion segmentation becomes difficult due to the presence of low contrast, irregularities, rough boundaries and different artificial factors [164, 165], which have been mentioned in section 3. Generally, before inputting the image into the segmentation algorithm, effective image preprocessing should be used to eliminate the influence of these factors [166].

5.1. Lesion Boundary Detection

In the actual clinical diagnosis environment, boundary labeling is performed handcrafted by doctors. This is a tedious step, and its results are easily affected by the subjective nature of the doctor, which will affect the final clinical results. Therefore, a necessary task of computer-aided diagnosis is the boundary detection of lesions, which is also used as a preprocessing step in many literatures. Automatic detection of skin damage boundary can bring better segmentation effect of skin slice image, effectively increase the efficiency of skin damage recognition, and help health care providers take care of patients in a better way.

Ideally, there is a smooth transition between the lesion and the surrounding normal area, and the boundary is a group of continuous pixels in the smooth transition. The basic idea of boundary detection is to find the position where the intensity changes rapidly in the image, usually where the first derivative is greater than the specified threshold or where the second derivative crosses the zero point [167].

Traditional boundary detection techniques are usually based on threshold, clustering, region growth and so on. However, due to the low contrast between the surrounding skin and the lesion, the blurring of the lesion boundary, artificial artifacts such as hair, and color changes inside the lesion, it is a challenge to automatically detect this boundary. Ali *et al*. [168] proposed a new edge-extracted approach for skin lesion boundary detection. This method could detect the main boundary around skin lesions and be robust to artifacts presented in the image. This had good applicability in the preprocessing step of the lesion segmentation task. However, experiments showed that this method lacked the ability to detect the fine structure of the skin lesion boundary. The hyperparameters introduced by traditional image processing techniques hinder irregular contour detection. Jayalakshmi *et al*. [169] utilized a median filter to remove artifacts and proposed an improved K-means clustering method. It was verified on the public Danderm database that the clustering algorithm performed well in detecting lesion boundaries and was suitable for preprocessing of dermoscopy images. Contour refinement is a difficult task in lesion edge detection. Complex lesions may hinder the refinement performance of the model. After preprocessing to improve the image quality of skin lesions, Sengupta *et al*. [170] utilized an ant colony optimization algorithm to improve the boundary contours of skin lesion images processed by conventional boundary detection methods. The introduction of this method improved the efficiency of conventional boundary detection methods in skin lesion images. The complementarity of models and strategies enables the combinatorial optimization between them to deal with more complex focal areas. Abbas *et al*. [171] tried to combine different methods to improve segmentation performance, among which the combination of bilateral filter, polynomial model and Canny boundary detector could achieve the best average accuracy. Bayraktar *et al*. [172] used a new local boundary extraction method and probability map to overcome the problem of boundary blur in active contours, and the effect on dermoscopy images was not inferior to the latest methods. Abeysinghe *et al*. [173] proposed two new methods (distance difference method and gradient method) with different ideas from existing methods to detect border irregularities, thereby making medical image detection easier.

In the study of clinical image data, due to the unlimited shooting distance, a single image sometimes contains multiple lesion areas and a large area of skin. Compared with normal skin, the skin lesion area is sometimes very small in the image, so a certain lesion area detection is needed. The image data available for training specific diseases is generally insufficient. Therefore, Han *et al*. [174] adopted the R-CNN model to detect all small lesions from a limited number of large-size images, thereby solving the problem of insufficient clinical image datasets. Furthermore, the use of neural networks to predict the irregularity of the skin lesion boundary has also become an additional application of boundary detection. Ali *et al*. [175] proposed a method for determining whether the boundaries of skin lesions were irregular. The surface irregularity of the skin lesion detected by It was used as the input of CNN, and Gaussian Naive Bayes was used to detect the irregularity of the boundary objectively and automatically. In addition, Ali *et al*. [176] also designed a new fuzzy multilayer perceptron and corresponding activation function. The proposed method generally outperformed most of the latest standard neural network classification methods.

5.2. Semantic Segmentation

The semantic segmentation task is also the basic task of computer vision, which is essentially to classify every pixel in the image at the semantic level, and the result will have strong real semantic interpretability. At present, medical image segmentation is one of the main fields involving semantic segmentation tasks. Semantic segmentation, compared to classification task, there are a lot more challenging problems: Firstly, the classification task only requires that level of output image category, and the output of the semantic segmentation task not only needs to include the category, the pixel location information also need not precise output, whether wrong category or wrong location information can lead to bad model performance; Secondly, the semantic segmentation task involves the output of every pixel in the image, no matter in the training stage or the testing stage. Compared with the classification task, which contains only one image-level output, it undoubtedly carries heavy computational cost. Finally, context information plays an important role in segmentation task. Most classification networks reduce feature dimension to obtain accurate model effect, but it often leads to irreversible information loss.

The goal of semantic segmentation task is to segment the fine structure under the large background, which has important research significance in the application of computer-aided diagnosis systems. In recent years, semantic segmentation in medical context has been widely studied, such as multi-target labeling based on the living microscope. When a dermatologist analyzes the histopathological features of the skin, individual histopathological features of the skin are associated only with certain local areas in the image. However, medical images often have extremely complex, noisy structures. For example, there may be multiple histopathological features. At this point, if there is a method to divide the image into multiple regions and extract each region structurally, it can lay a good diagnostic environment for deeper disease classification.

When a dermatologist analyzes the histopathological features of the skin, a single histopathological feature of the skin is only related to certain local areas in the image, and the image may have multiple histopathological features. At this time, it is necessary to divide an image into multiple regions, and perform feature extraction on each region, to lay a good diagnosis environment for deeper disease classification. RCM images are very different from other images in imaging principle, Zhang *et al*. [177] demonstrated a feature representation method for skin biopsy tissue pathological image annotation based on deep learning. Bozkurt *et al*. [178] described a new multi-resolution convolutional network structure and used it to annotate human bodies. Morphological patterns in skin RCM images. Goyal *et al*. [179] proposed a multi-class segmentation method based on FCN to segment dermoscopy images of benign moles, melanoma, and seborrheic keratosis. However, based on dermoscopy images and clinical images, it is often necessary to complete the task of segmentation of skin lesions and background. On the high-resolution clinical images of multi-focal areas, Liu *et al*. [180] focused on the four types of segmentation problems of cutaneous T-cell lymphoma (CTCL) and similar diseases. A novel multi-knowledge learning network was proposed to solve this problem, which achieved very good performance and met clinical needs.

In addition to deep learning methods can achieve multi-class segmentation tasks of diseases, some non-deep learning algorithms can also solve part of semantic segmentation problems [181]. In an automatic framework, dictionary learning [182] and graph-cut [183] are combined to achieve a high level of multi-class segmentation performance. By constructing the image input dictionary and updating the corresponding label dictionary, the feature data set with optimal value is obtained. Finally, the graph-cut method with label cost is used to obtain better results than the most advanced methods.

However, when a deep learning model is used for end-to-end training, although excellent performance can be obtained, the abstract features are not interpretable at all, which will greatly hinder the application of model algorithm. Thomas *et al*. [184] provide a new deep learning model interpretability method, which uses a semantic segmentation model to segment multiple types of pathological structures based on histopathological maps of common skin cancer types. It should be noted that the semantically segmented categories correspond to the high-level concepts in human cognition, providing an interpretable clinical application mode for the future computer-aided diagnosis system.

5.3. Segmentation Algorithm Based on Traditional Machine Learning

The handcrafted features of machine learning include color features, texture features, and shape features. Combining these handcrafted features, threshold [185], region [186], and morphology [187, 188], these methods have been studied to prove that they can provide better for specific tasks. The result of the diagnosis. Patiño *et al*. [189] used the SLIC algorithm to segment the image and combined the average RGB colors of superpixels to effectively deal with hair removal, oil bubbles, light changes, and reflection images. In order to reduce the impact of low-contrast boundaries on clinical image segmentation, Filali *et al*. [190] started from the image super-pixel level and used simple linear iterative clustering (SLIC) and image propagation for contrast refinement, surpassing some of the most advanced method on the two datasets. However, the disadvantage of the method based on color space and threshold is that it can only process color-based features with the same size, has a high dependency on the threshold, and lacks generalization ability. Devi *et al*. [191] proposed automatic cluster selection using Fuzzy C-means based on histogram attributes. The system segmented melanoma from non-dermoscopy images of normal skin. Salih *et al*. [187] used random region fusion combined with a pixel-based Markov random field model to achieve skin lesion segmentation.

In addition, the method based on texture analysis can effectively analyze all texture features. Peruch *et al*. [192] utilized Markov random fields (MRF) and principal component analysis (PCA) to obtain good results. Ma *et al*. [193] went beyond the limitations of traditional directions and used deformable models for model segmentation of curve evolution. Pereira *et al*. [194] used LBP and gradient-based histogram thresholding (GHT) methods to extract features, and utilized SVM to get good segmentation results. Hasan *et al*. [195] proposed a new Segmentation-based Fractal Texture Analysis (SFTA) method for texture feature extraction. This method combined a hybrid multi-level threshold algorithm to select the optimal threshold number, which was 2% more accurate than the traditional Otsu threshold-based SFTA.

Researchers focus on dermoscopy images, and introduce local and global features based on handcrafted features. Slowly some researchers began to think about whether combining a single method could bring better segmentation results. Priyadarsan *et al*. [196] used the combination of local variance and global threshold segmentation method had obtained higher accuracy than existing segmentation methods. Global features bring new information to the model, and feature information of different dimensions brings information to the model correlation. Ruela *et al*. [197] suggested that global and local features were based on two types of fusion (early and late fusion) technology. Nasir *et al*. [198] proposed a hybrid method that extracts HOG (shape), color, and texture features and used the additive law of probability to achieve the best results on the PH2 dataset. Asaeikheybari *et al*. [199] designed a Multiple Random Walker segmentation algorithm and compared it with three CNNs used for segmentation to verify the diagnostic effect of the Multiple Random Walker algorithm structure in the case of limited datasets.

Spectral analysis of histopathology brings positive enlightenment to skin tumor representation. McIntosh *et al*. [200] considered the fact that infrared light was absorbed by a variety of skin components, and studied the application of infrared spectroscopy in the characterization of basal cell carcinoma specimens *in vitro*. Furthermore, linear discriminant analysis (LDA) was used to analyze the near-infrared absorption spectra for the non-invasive *in vivo* characterization of skin tumors [201].

5.4. Segmentation Algorithm Based on Deep Learning

In order to reduce the misjudgment rate of the algorithm, it is often necessary to learn the features that maximize. In the research process of several machine learning algorithms based on artificial features, the existing methods still cannot fully learn the effective features for diagnosis, nor can they detect the precise division of the boundary area. Although the deep learning model is very complex, it can learn deep heterogeneous features from the original image, and can express different levels of information from traditional handcrafted features. Among them, CNNs have excellent effects in the field of image processing, especially medical images.

Skin lesion image segmentation is a difficult task in computer vision. Deep learning techniques, especially convolutional neural networks, have achieved great success in this regard [65, 75, 89, 202-204]. Researchers have proposed various models based on deep learning. Some structures including U-Net, Fully Convolutional Network (FCN), Fully Convolutional Residual Network (FCRN), Convolutional Deconvolutional Neural Network (CDCNN) and GAN) have all produced skin lesion segmentation Excellent performance. FCN is integrated by the convolutional layer and the pooling layer, and is one of the first models proposed for segmentation. Kaymak *et al*. [205] compared the effects of four different FCN architectures on the public dataset. However, the FCN model has the possibility of over-segmentation, which may cause the segmentation effect to not reach refinement [206]. The U-Net architecture is developed from FCN. It is an encoder-decoder network, where the encoder and decoder parts are connected by shortcut skip. But its weakness is that it will cause loss of information during short skip, resulting in incomplete inclusion of image features in the decoding process. Considering the excellent performance of ResNet [207] and DenseNet [208] in image classification tasks, people combined the idea of residual blocks or dense blocks into the existing image segmentation architecture, and designed related models [209]. The more common one is FCRN. But running FCRN-based architecture requires a lot of computing resources, which may limit the use of the architecture in actual scenarios. The architecture of CDCNN is composed of convolution and deconvolution networks. The deconvolution layer is used to smooth the segmentation image in order to obtain the final high-resolution output. However, the implementation of this architecture also requires high computational costs. Thanks to the great success of GANs [95] in image generation tasks, the idea of adversarial training has been used by people to construct effective lesion segmentation networks, and gratifying results have been achieved [204, 210-214].

The model of the segmentation task includes down-sampling and up-sampling parts, where the down-sampling part essentially obtains a feature extractor. Therefore, for a very limited number of medical images, the segmentation task is often initialized with the pre-trained parameters of the classification model and fine-tuned. Tschandl *et al*. [215] trained VGG and ResNet classification networks on public datasets, and then transferred the corresponding layers as encoders to the LinkNet model [216] and fine-tune them. Compared with the randomly initialized network, a model with fine-tuned weights achieved a higher Jaccard index at public dataset. Soudani *et al*. [217] used two pre-trained networks and then considered building a five-node classifier to predict the most suitable segmentation technique. Phillips *et al*. [218] proposed the idea of preprocessing multiple pre-trained networks on the PascalVOC [219] segmentation dataset and fine-tuning in multiple steps on the entire training set image. This model achieved the effect of handcrafted fine segmentation. The weights of the existing pre-trained networks were all trained on ImageNet or PascalVOC, but the visual features of the images in these datasets were more shapes, which were not applicable under the irregular conditions in the field of skin diseases. In order to solve the induced feature deviation of the pre-trained dataset, Canalini *et al*. [220] explored three pre-trained strategies based on the segmented structure to initialize the feature extractor, but different pre-trained networks focused on different Features, thereby greatly improving the effectiveness of the integration. The detection task model can be roughly divided into one-stage method and two-stage method, and the combination results of different methods can improve the model performance. This is because different models approximate the final model results through different principles and mechanisms, and the appropriate result combination strategy can produce synergies between models. Bagheri *et al*. [221] verified that the two-stage combination strategy could more effectively combine the output of Mask R-CNN and Retina-Deeplab model. The specific approach is to first do post-processing to the output of the two models to get the segmentation result, and then do post-processing to the segmentation result to get the final more accurate segmentation model. Hasan *et al*. [222] enhanced the result of lesion identification by integrating multiple convolutional neural network feature extractor modules. Finally, an automated dermoscopy image diagnosis Web application was designed and validated on an ISIC dataset. Xiao *et al*. [223] designed a few-shot learning method for the lesion segmentation task, considering the high cost of obtaining tags for the lesion segmentation task. Symmetric networks were proposed to perform repeated training networks for the same set of data under the premise that most of the background was removed by using the common regions of the support set and validation set data. The evaluation results on open data sets show that few-shot learning networks based on a small amount of annotated data have good prospects. Facing the same situation, Jin team [224] obtained time-saving automatic weak labeling of data by using the threshold method, which can obtain a large number of weak labeling data in a limited time. Through experimental verification, weak labels of automatically generated data can be used for supervised training of deep learning, and obtain 78% accuracy. Some unsupervised segmentation algorithms were designed to save a lot of data labeling costs and label bias. Messadi *et al*. [225] constructed an unsupervised segmentation algorithm for lesions, which reduced 6% compared with Growcut and Mean Shift under the edge error index.

Compared with traditional machine learning, deep learning models can learn more comprehensive and deep data features, and may achieve better results on certain tasks. Lin *et al*. [226] compared U-Net and C-means methods on the task of skin lesion segmentation, and the results showed that compared with clustering methods, U-Net method had obvious advantages. The preprocessing process of dermatological images is very challenging, and the combination of deep learning and traditional machine learning methods sometimes fits well. Traditional machine learning algorithms show good performance in dataset preprocessing, and can be used as a good auxiliary method for deep learning before final segmentation. Huang *et al*. [227] used K-means clustering and improved Mask R-CNN to achieve good segmentation results in segmenting skin lesions, and solved the problem of fuzzy boundaries and complex textures. Morphological algorithms can deal with artificial artifacts better because of their excellent adaptability. Justin *et al*. [228] used morphology-based hat transformation to preprocess the image, and then adopted DeeplabV3+ to achieve efficient lesion segmentation. Zafar *et al*. [229] used a hair removal algorithm to preprocess the image, and achieved an effect that was not inferior to the advanced level on the improved U-Net. Li's team [230] designed a new idea for hair preprocessing, which is based on U-Net deep learning model to evaluate the hair removal effect of a single image, and achieved better results than other advanced algorithms on the ISIC2018 dataset. Ramya *et al*. [231] preprocessing images of skin lesions based on discrete wavelet decomposition of different color components, and obtained clear segmentation masks of lesion regions by threshold method after processing the complexity of images. The complexity of the lesions is significantly different from normal skin color, but is often limited by the Angle of shooting and the influence of light in the study. Dastane *et al*. [232] used two-stage pixel neighborhood technology to realize discrimination. Specifically, the classification recognition probability of each pixel is obtained by using the deep learning model, and then the original skin color information is better utilized by combining the auxiliary information of pixel probability in the neighborhood. The Filali team [233] used a graph-weighting method to control the relative weight of feature areas near the edges so that the model could more easily adapt to light noise in the macro image. Compared with other segmentation methods, the result of this method is more accurate and faster. In contrast, compared with CNN's convolutional operation, which constantly sacrifices image resolution to increase local receptive fields, some machine learning strategies are more advantageous in the refinement requirements of segmentation results. Adegun *et al*. [234] adopted CNN as the initial segmentation method of the lesion, and then performed boundary fine segmentation with the watershed algorithm. In order to build the number of parameters of the model and extract relevant features, Xu *et al*. [235] optimized CNN by satin bowerbird optimization (SBO). This algorithm always aims to find a parameter combination optimization. Similarly, the World Cup Optimization algorithm and Imperialist Competitive Algorithm were used to refine the boundary of lesions [236, 237]Ünver *et al*. [75] combined the YOLO model to locate the lesion to refine the segmentation effect of the GrabCut segmentation algorithm, which greatly improved the evaluation index of the original algorithm. Adegun *et al*. [238] used a probabilistic model based on conditional random fields to refine the boundary of the output results of the FCN model. The system uses a lightweight design and achieves better performance on two open data sets. Qiu *et al*. [239] maximized the label consistency between similar pixels through condition random field (CRF), to achieve the goal of refinement and coarsening of the pixel prediction of multiple deep CNN models to generate fine-grained segmentation. Neural networks usually consider resizing the original image to save computational cost, but the behavior of reducing the resolution may cause information loss. Masni *et al*. [89] developed a skin lesion segmentation method through a deep full-resolution convolutional network (FrCN). This method could accept the original resolution of the input image without pre-processing or post-processing resize operation transformation. Sometimes the segmentation task of melanoma based on deep learning features cannot achieve high accuracy, and the output abstract features are easily misled by false features. Khan *et al*. [240] extracted depth features and combined an improved moth Flame optimization algorithm to remove irrelevant and redundant depth features, achieving significant improvement results.

Based on the infrastructure proposed in the existing research literature, many studies modify existing networks and expand them to be more suitable for specific tasks. Shan *et al*. [241] referred to the DPN network structure to improve the block structure in FC-DenseNet, and set up comparison experiments with three different segmentation models. The global context can correctly guide the performance of the model, thus correcting the wrong learning direction of the model in time. U-Net has great advantages over other segmentation models in combining context information. Therefore, Jiang *et al*. [242] added a consistency monitoring module of global context before sampling input on U-Net, which has better performance than the current best methods. Common splicing operations in the U-Net model will undoubtedly lose some context information. Qamar team [243] designed a dense jump module, which uses the jump idea of dense network and different expansion rates of pyramid pooling [244] to capture global context information. The model based on this module has achieved the most advanced performance. The results showed that the improved model maintained a tradeoff between model complexity and overall performance. The attention module has been widely used in recent research due to its excellent performance. People combine the attention module into the existing image segmentation architecture to design an effective deep network for skin lesion segmentation. Xie *et al*. [245] designed a high-resolution attention function block with three branches. By fusing spatial attention and channel branch output, robust features with detailed spatial information could be extracted. In this way, proposed method could obtain accurate skin lesions boundary. Sarker *et al*. [246] combined the lightweight GAN model with the position and channel attention module. In the case of reducing the amount of model parameters, considerable performance was still obtained on the public dataset. Similarly, Jiang *et al*. [247] proposed a CSARM module, which combined residual learning, channel attention mechanism and spatial attention mechanism to improve the discrimination and representation capabilities of CNN, and compared other attention modules on the public dataset, such as SE Block and FPA. Tao *et al*. [248] constructed a channel attention module that can capture multi-scale features to improve the performance of the model on the open dataset HAM10000. Tong *et al*. [249] added three different forms of attention mechanism to the classic U-Net model: channel attention, spatial attention and contextual attention controlled by the gate. The experiment proves that such attention modules can indeed bring more relevant visual attention to the target the network. Arora *et al*. [250] proposed an attentional gate to capture high-dimensional features from low-dimensional irrelevant features in the skin segmentation task, and evaluated it on ISIC2018 dataset. As attention modules are more frequently used in deep learning models, Ren *et al*. [251] 's work involves the effects of different combinations of different types of attention modules. The research object is the channel attention module and the spatial attention module, the main research combination is the number, sequence and combination mode. This study verified that the serial combination of channel attention and spatial attention was more conducive to the aggregation of global and local information in the segmentation task, and finally achieved an average Jaccard index of 0.7692 on the ISIC2017 dataset.

Different information features in medical images may all play important roles in specific tasks. In addition, the same feature in different tasks will also be emphasized due to actual index requirements. In deep learning segmentation models, deep feature maps are often used as input for upsampling. Unlike deep features, which are generally abstract features, shallow features are more inclined to color, shape, and location. Therefore, the literature is often willing to combine low-level handcrafted features, hoping to provide guidance for diagnosis from an explanatory perspective. Color and texture features are the most important attributes of dermoscopy images, and these are one of the important features used to identify skin diseases [253]. Although handcrafted features usually lack generalization ability, they show poor performance compared with deep neural network features learned directly from large amounts of data. However, they can sometimes achieve excellent performance in certain dermatological diagnosis tasks, which can be used as a supplement to deep features. Researchers began to combine handcrafted features on the original neural network structure to improve the performance and interpretability of the model. The deep features extracted by deep learning models are always lack of reasonable interpretability, which will be seriously questioned in the actual implementation of medical models and is not conducive to the study of medical symptoms. Reasonable feature selection methods can extract the depth features related to disease categories, so as to provide some suggestions for practical medical research significance. Kaya *et al*. [253] utilized parameters and non-parametric correlation coefficients to conduct correlation ranking for deep features, so as to eliminate redundant features that are relatively unrelated and achieve better model performance after eliminating redundant features. Messadi *et al*. [225] effectively combined traditional features (ABCD rules) on an artificial neural network to achieve an increase in the rate of melanoma recognition of true positives, while ensuring the interpretability of the depth model. Codella *et al*. [252] proposed a fully convolutional U-Net structure that combines RGB and HSV channel input, which could better explain the feasibility of diagnosis that met the understanding of physicians. With continuous breakthroughs, Yuan *et al*. [254] expanded the early convolutional-deconvolutional network (CDNN) model and combined multiple color channels as a dual threshold mask to achieve better segmentation results. Considering that the lesion area generally had obvious texture deformation, Kaur *et al*. [255] introduced a hybrid deep learning method that used feature vectors based on traditional textures as input to train deep neural networks. It achieved excellent performance in RCM skin disease image recognition. Pour *et al*. [256] verified that images from the transform domain had the potential to improve performance on the improved model of the CIELAB color space. Abhishek *et al*. [257] proposed a depth segmentation framework that used additional color channels and light-invariant intrinsic gray and shadow attenuation images to enhance RGB dermoscopy images, and evaluated the proposed method on three datasets Effectiveness. Different models have different emphases on learning objectives, so model integration can bring new enlightenment to both deep learning models and traditional machine learning algorithms. Khatibi *et al*. [259] firstly input the pre-processing images into four deep learning models to extract abstract features, and then use three unsupervised machine learning algorithms to aggregate the segmentation results of the abstract features after spliced. Experimental results show that 97% of the segmentation accuracy is achieved on 877 large resolution data sets.

In some studies, the idea of classification models will also be used to improve performance by combining multiple models. Due to the difference in network depth and module structure, different models will have different emphases on the same data set or task learning. Attia *et al*. [258] proposed a combination of FCN and LSTM [260] to segment melanoma images, and concluded that the hybrid method of RNN and CNN was superior to the method that only relies on CNN. The combination of different levels of features in the model training process to obtain more matching information has also aroused strong interest among scholars. Bi *et al*. [261] proposed multiple embedded FCN stages to learn different levels of visual features of skin damage, and fused these features together to accurately segment skin damage. Li *et al*. [262, 263] performed hierarchical supervision to obtain low-level boundary information and used chain residuals to fuse multi-level features. Ji *et al*. [264] used the supervised block to learn the output features in the up-sampling stage of the modified U-Net model, and finally integrated the multi-path output to obtain better performance. Liu *et al*. [265] Introduced distance metric-based learning before the input of the classifier, and obtained the intermediate feature representation by using the relationship between different image samples, so that performance will be more robust when dealing with large intra-class differences and inter-class similarities. Bozorgtabar *et al*. [266] designed the side output to explore the role of the features in the middle layer. In the end, in addition to outputting the segmentation contour probability map, the fuzzy boundary of the lesion can also be obtained to provide visualization. Nathan *et al*. [267] introduced coordinate convolution before passing the input image to the encoder. This helped the network to determine features related to translation invariance which further improved the generalization ability of the model. The main advantage of multi-scale is to provide accurate boundaries with different scales and contrasts, which brings great help to skin disease diagnosis tasks [268]. Singh *et al*. [269] used a multi-scale input strategy to select filters with variable scales, which better matched specific skin segmentation tasks. Zhu *et al*. [270] presented a novel adaptive scale module, which could effectively dynamically integrate multiple scale information and provide greater self-learning ability. Bi *et al*. [271] adopted a multi-scale strategy to scale the image input to seven scales, and performed a flip operation on each scale to segment skin lesions, model with these operations successfully improved the over-segmentation or under-segmentation. Jafari *et al*. [58] simultaneously used the local and global regional features of clinical images to make decisions on the final output after learning the local and global information independently. Although the prior knowledge of the target object has been proven effective in skin disease diagnosis for a long time, few have embedded prior knowledge into a deep learning framework. Mirikharaji *et al*. [272] encoded the shape of the lesion as prior knowledge of FCN. By further penalizing counterexamples in the loss function, it could bring a good improvement based on different original segmentation models.

In addition to changing the novel model, people are also considering the development of effective deep learning models for skin lesion segmentation from other aspects. Goceri *et al*. [273] designed a new adaptive and asymmetric loss function and verified its superiority for the task of lesion segmentation on ten networks. Zhang *et al*. [274] proposed a new loss function based on Kappa index, which could be used for medical image segmentation in CNN. Unlike Dice loss, this loss function considers all pixels (including background pixels) in the evaluation of predictive segmentation. In some cases, Kappa loss helped to make segmentation more accurate on six datasets. The image segmentation task itself has a large area of background area, so it is easy to produce the phenomenon of category imbalance in pixel-level segmentation task. Similarly, the loss of Dice directly ignores the punishment of negative sample pixels, which will greatly mislead the training of the model. Abhishek *et al*. [275] designed Matthews Correlation Coefficient Loss to monitor the confidence of positive and negative samples at the same time, and verified the conclusion on open data sets. Hasan *et al*. [276] proposed a hybrid loss function on a new lightweight segmentation model, which maximized the prediction area to approximate the true value. Zhang *et al*. [277] adopted an improved optimization algorithm to optimally select the weights and deviations in the network, and optimized the efficiency results of CNN. At the same time, for skin damage segmentation, some people find that a considerable part of the dataset had a low degree of consistency with the true value, which indicated that there would be a certain difference in training. Ribeiro *et al*. [278] removed noisy samples from the dataset, and removed excessive details from the boundary truth values of the remaining samples to improve the effect. Since the method of the deep segmentation model cannot be easily extended to datasets with multiple image annotations, Mirikharaji *et al*. [279] suggested an ensemble learning scheme to effectively deal with the differences in segmentation annotations. This method improved the generalization performance of the deep segmentation model by using all available annotations.

As more and more excellent deep learning models are designed, the types of medical images studied are almost all concentrated on the dermoscopy image dataset, The publicly available large-scale dermoscopy dataset has excellent imaging quality and high field recognition. Although the clinical images that are easier to obtain in comparison have better appearance geometric features, clinical images are rarely used in skin segmentation tasks. Because the observation range of the original clinical image is too large, it is a big challenge for the algorithm to learn the characteristics of the lesion area. Raj *et al*. [280] segmented psoriasis lesions in clinical images with complex backgrounds and challenging environments. Udrea *et al*. [281] obtained segmentation of lesions of images acquired by mobile devices based on deep networks of GANs. Based on many images acquired by smartphone cameras, the performance of the network was verified. Lesion segmentation has also been tried on more types of datasets and disease types. Biopsy images contain more internal information about the lesions. For this type of images, it is often necessary to consider preprocessing artifacts caused by staining. Pal *et al*. [282, 21] studied the use of deep learning methods to achieve this. A segmentation task of small biopsy tissue image datasets for psoriasis. Bozkurt *et al*. [283] conducted an in-depth study on the structure of deep neural networks applied to RCM image classification, and proved that this task had a significant improvement over the previous latest results[284]. Deep learning medical image segmentation model can also be applied to high-frequency ultrasound (HFUS) images. Czajkowska's team [285] used high-frequency ultrasound (HFUS) images to automatically segment skin layers, outperforming current state-of-the-art algorithms on 380 patients. The practical application of the algorithm often meets many constraints, among which the dermoscopy acquisition equipment obviously does not have a high-performance processor and high-capacity memory. Sarker's team [286] came up with a new lightweight GAN model to accommodate this deployment environment. One-dimensional kernel decomposition is used to reduce the burden of two-dimensional convolution filtering, and multi-scale feature fusion, channel and spatial attention monitoring and binary loss function are used to improve the performance of the lightweight model. Based on the U-Net model, Wibowo *et al*. [287] combined LSTM and depthwise separable convolution structures, respectively, and compared MobileNet V3-UNet, a lightweight, high-performance model that was more suitable for the task of segmentation of lesions, with better performance than several advanced methods on three public data sets.

In order to facilitate reference to the current literature on the segmentation of skin lesions, we list Table **3** below by year.

Table 3. Current research reference of skin lesion segmentation.

Refs.	Year	Dataset (Category)	Model & Method	Remarks
[202]	2001	195 cases (5)	LDA	Using uniform segmentation and feature selection based approach
[193]	2013	30 images (2)	PCA + MRF	SLIC and image propagation for contrast refinement
[184]	2015	250 images (2)	CIELAB + Geometric model	Automatic cluster selection using Fuzzy C-means based on histogram attributes
[196]	2016	KTH-TIPS (10)	SFTA + Threshold	Deformable models for model segmentation of curve evolution
[257]	2016	1,500 images	Hybrid CNNs	Improving the block structure in FC-DenseNet by referred to the DPN network
[274]	2016	Dermquest	CNNs	Images from the transform domain
[180]	2017	ISIC-2017	FCN-8s	High-Resolution Clinical Images
[198]	2017	169 images (2)	HSV + KNN	New Segmentation-based SFTA method for texture feature extraction
[228]	2017	ISIC 2017	U-nets + C-Means	Few-shot learning method for the lesion segmentation task
[254]	2017	ISIC 2016	U-net + RGB&HSV	Maximizing the label consistency between similar pixels through CRF
[256]	2017	ISIC 2017	FCNs	Depth features, combining improved moth Flame optimization algorithm
[260]	2017	ISIC 2016	RNN + LSTM	Atrous convolution
[263]	2017	ISIC 2016	FCNs	Lightweight generative adversarial network
[268]	2017	ISIC 2016	FCNs	Serial attention network
[284]	2017	3,000 images	GAN	Coordinate Convolution and Deep Residual Units
[287]	2017	504 images	CNNs	Using a multi-scale input strategy to select filters with variable scales
[178]	2018	12,600 images (4)	CNN	A multiresolution convolutional neural network with partial label training
[179]	2018	56 images (4)	U-net	A multi-class segmentation method based on FCN
[190]	2018	PH2	SLIC	Using histogram thresholding on optimal color channels
[199]	2018	PH2	Boltzman Entropy + SVM	The combination of local variance and global threshold segmentation method
[264]	2018	ISIC 2017, PH2	FCN	Digital hair removal by deep learning
[265]	2018	ISIC 2016, 2017	CNN	Basing on Multi-Scale Attention Convolutional Neural Network
[266]	2018	ISIC 2018	U-net	Attention Gate, Spatial and Channel Attention U-Net
[285]	2018	90 images	U-net	Coordinate convolution before passing the input image to the encoder
[191]	2019	206 images (2)	SLIC	Random region fusion combined with a pixel-based Markov random field model

Table (3) contd….

Refs.	Year	Dataset (Category)	Model & Method	Remarks
[200]	2019	PH2	Multiple Random Walker	Utilizing shape and symmetry features
[217]	2019	ISIC 2017	LinkNet-152	Using two pre-trained networks and a five-node classifier
[219]	2019	ISIC 2017	ResNet-50	Preprocessing multiple pre-trained networks
[220]	2019	50 WSIs	FCNs	Three segmented structure pre-trained strategies initialized the feature extractor
[222]	2019	ISIC 2018	DeepLabv3+ + ResNet-101	Integrating multiple convolutional neural network feature extractor modules
[248]	2019	PH2, ISIC 2017	Yolov3 + GrabCut	Morphology-based hat transformation and DeeplabV3+
[250]	2019	ISIC 2017, 2018	GAN + Pix2pix + Attention	A graph-weighting method to control the relative weight of feature areas
[270]	2019	ISIC 2016, 2017	GAN	A fully convolutional U-Net structure that combines RGB and HSV channel input
[181]	2020	57 images (4)	Encoder-decoder	Jointing dictionary learning
[187]	2020	170 images (2)	Fuzzy C-Means	Image bit-plane multilayer approach
[188]	2020	PH2, ISIC 2018	MRF	The SLIC algorithm combined the average RGB colors of superpixels
[192]	2020	MED-NODE, Dermofit	GHT+ LBP+ SVM	Using MRF and PCA
[195]	2020	ISIC 2017, DermQuest	Threshold + morphological	LBP and GHT to extract features; SVM to get good segmentation results
[206]	2020	ISIC 2017	FCN-8s	A convolutional neural network with an attention mechanism
[229]	2020	23,906 images	CNN + K-Means	Automatic weak labeling of data by using threshold method
[230]	2020	PH2	Deeplab V3+	An unsupervised segmentation algorithm for lesions
[231]	2020	ISIC 2017, PH2	U-net + ResNet	U-Nets versus clustering
[236]	2020	PH2	U-net	K-means clustering and improved Mask R-CNN
[237]	2020	ISIC 2017, PH2	CRF + 15 CNNs	The hair removal algorithm to preprocess the image
[258]	2020	ISIC 2017, PH2	FCN + DPN	Discrete wavelet transform
[259]	2020	ISIC 2017, 2016, PH2	CNN +Attention	Two-stage pixel neighborhood technology
[267]	2020	ISIC 2017, PH2	U-net + Attention	Performing boundary fine segmentation with the watershed algorithm
[269]	2020	ISIC 2017	U-net + CIELAB	Consistency monitoring module of global context before sampling input on U-Net
[272]	2020	ISIC 2017, DermoFit, PH2	U-net	A dense jump module
[277]	2020	ISIC 2017	ResNet-50+DenseNet-201	Attention-based deep convolutional neural network
[279]	2020	ISIC 2016, 2017, ISIC 2018, PH2	U-net	Parameters and non-parametric correlation coefficients to conduct correlation ranking
[281]	2020	ISIC 2018	ResNet-34	Combining multiple color channels as a dual threshold mask
[283]	2020	ISIC 2016, 2017, ISIC 2018	U-net	Feature vectors based on traditional textures as input
[293]	2020	ISIC 2017, DermoFit,, PH2	U-net	Multistage fully convolutional networks
[294]	2020	ISIC 2017, PH2	U-net + FCNs	Multiple embedded FCN stages and fused features together
[295]	2020	Dermquest, DermIS	CNNs	Dense connected deconvolutional network
[296]	2020	ISIC Archive	DeepLab V3+	Hierarchical supervision and using chain residuals to fuse multi-level features
[297]	2020	ISIC Archive, PH2, DermoFit	FCNs	Middle-level feature learning
[298]	2020	350 images	U-net	Distance metric-based learning before the input of the classifier

6. MULTI-TASK MODEL

For the tasks in the skin disease image intelligent analysis, recent scholars have done a lot of research work. For example, from the perspective of image preprocessing, model structure, diagnosis algorithm, loss function, etc., to improve the diagnosis effect, or to promote the application of the model from the perspective of the applicability of the task and the limitation of the disease. But most of the research goals are a single task, strictly speaking, this is a limitation for the medical diagnosis task itself.

6.1. Multi-task Model Based on Structure

CAD has become an important research hotspot of artificial intelligence in the medical image analysis field. CAD usually includes five steps: data collection, data preprocessing, lesion segmentation, feature extraction, and diagnosis and classification. The above contents are all people's research and elaboration on the single step in CAD. However, in skin disease diagnosis, sometimes it is often necessary to develop a two-stage framework. The segmentation of the lesion is input to the classification model to achieve higher accuracy. The segmentation task is used to detect the location and boundary of the lesion and extract the lesion area, while the classification task is used to diagnose the type of lesion. Nevertheless, the segmentation and classification of skin lesions are still two highly related tasks. Skin lesion segmentation helps to remove interference in image data (such as dermoscopy images) and improves the accuracy of skin lesion classification, while category-specific diagnostic information also helps to highlight the area of the skin lesion, thereby helping the skin lesion Segmentation.

Some studies have begun to propose that the output of simple segmentation is used as a model for classification tasks, and the diagnosis results can be significantly improved. Some studies have implemented a two-stage model using traditional machine learning methods. According to the feature information of the dark spots or droplets on the segmented images, Maglogiannis *et al*. [288] implemented an efficient two-class classification of melanoma through a classifier. Premaladha *et al*. [289] utilized threshold segmentation in combination with classifiers in benign and malignant tasks to overcome the light and shadow interference of the image and achieved high accuracy. Patiño *et al*. [290] first adopted the super pixel merging strategy with RGB criteria to segment the lesions, and realized three classifications by logistic regression combined with SVM. Deep learning algorithms have certain advantages in feature extraction, so there is research and application of deep learning models to achieve the stage of lesion classification. Aishwarya *et al*. [291] used K-means and Fuzzy C-means clustering algorithms to segment lesions, and achieved an efficient two-class classification of melanoma on the CNN model. Sikkandar *et al*. [272] developed a classification model in three steps, which were segmentation based on Grabcut algorithm, feature extraction based on Inception v4, and classification based on adaptive neuro-fuzzy classifiers. Goceri *et al*. [292] supplemented the automated detection of facial disorders (ADFD) segmentation algorithm with related denoising operations and proposed a new loss function. DenseNet201 was configured to achieve the best performance on the public dataset. Amin *et al*. [293] used wavelet transform and Otsu threshold algorithm to process segmentation tasks, and then used features extracted from two pretrained models of AlexNet and VGG-16 for classification. Compared with the results of existing work, this confirmed that the proposed method could classify skin lesions more accurately. Nazi *et al*. [294] compared nine different image enhancement methods on the original training images to improve the segmentation effect, and used DenseNet as a feature extractor to complete the classification of melanoma. Almaraz-Damian *et al*. [295] proposed a new CAD system to detect and classify malignant skin lesions. On the premise that the region of interest (ROI) was obtained, the fusion rules of interactive information were used to fuse the deep features extracted by CNN with the handcrafted feature ABCD rules related to medical algorithms, which verified the effectiveness of feature combination. Recently, there have been literatures that only use deep learning to achieve a complete two-stage task, which has gained some advantages over previous methods. Prathiba *et al*. [296] trained FCRN to extract the output of the lesion area from the skin disease image, and used these outputs for the residual network for melanoma classification. For realizing the final classification, Khan *et al*. [297] Input the R-CNN segmentation results into DenseNet, and adopted an entropy-controlled SVM to merge the two different levels of feature maps extracted from DenseNet. They discussed that in the future, feature vectors would be extracted on more architectures, and feature selection strategies would be improved to identify the most important features. Jayapriya *et al*. [298] finished a hybrid framework to combine two FCNs based on VGG-16 and GoogleNet, and used deep residual networks and hand-made features to extract features from segmented lesions to complete classification. Chang *et al*. [150] input the segmented image and the original dermoscopy image as a combination into the skin lesion classification network. The experimental results showed that the segmentation model and the classification model had achieved good performance on the international standard industry classification dataset. Han *et al*. [299] first used R-CNN to diagnose cancer by extracting lesion areas from more than 180,000 clinical photos. Experimental results showed that the accuracy of the algorithm was comparable to that of a dermatologist.

To compare the performance differences between the two assistive diagnostic processes, Maron *et al*. [300] trained two models for melanoma recognition on segmented and unsegmented dermoscopy images. Experimental results show that image segmentation plays an irreplaceable role in model recognition performance and can effectively remove the physical noise in the adjacent area of the lesion. However, in the actual execution steps, the performance of the whole model may be degraded if unqualified segmentation results are introduced. In order to further explore the effect of different input forms of skin lesion segmentation masks on the classification performance of dermoscopy images, Mahbod *et al*. [301] verified the results on the benchmark classification network, and adopted the segmentation mask in an appropriate way could significantly improve the overall classification performance. However, using the mask in an inappropriate manner by removing all background information greatly reduced the classification results. Some researchers have successfully verified that the segmentation task can help the performance of the classification network [302]. Conversely, the category positioning information of the classification network can also help achieve accurate segmentation. Xie *et al*. [303] proposed a deep convolutional network model that could be guided by each other to achieve both skin lesion segmentation and classification and diagnosis tasks. The ensemble network was mainly executed in two steps. The first was to use the mask generated by rough segmentation as the input of the classification network, which was beneficial to extract the effective features of the skin lesions. Then the skin lesion location information refined by the classification network was input into the enhanced segmentation to make the segmentation result more accurate.

6.2. Multi-task Model Based on the Loss Function

In the two-stage method, the model output of a single task is used as the prior work of the next task. Although satisfactory results can be obtained, but the knowledge learning process of the prior task is also worth exploring. Multi-task learning (MTL) aims to obtain better diagnostic features than the original task by sharing feature learning knowledge between related tasks.

Compared with the technology that solves a single task, the multi-task network technology is more robust and efficient. To a certain extent, it can effectively improve the learning efficiency and potential prediction accuracy of specific task models. In the field of assisted diagnosis of medical skin disease image analysis, researchers found the optimal performance of the model by sharing parameters among several parallel tasks. By constructing the task of combining different modal data, Kawahara *et al*. [156] also proposed to complete the classification of the seven-point checklist standard and skin disease diagnosis in one optimization. Pal *et al*. [17] regard the scoring tasks that affect the three parameters of psoriasis condition grading as interdependence and were designed in the same multi-task network. Using the classification of skin lesions as an auxiliary task, Liao *et al*. [304] trained a multi-task deep learning model. By jointly optimizing the two tasks of skin damage classification and body position classification, the performance of skin damage classification is significantly improved. In addition, there are studies to arrange several auxiliary tasks to improve the effectiveness of the target task. Vesal *et al*. [305] proposed adding a regional suggestion network to achieve target positioning, and used this as an auxiliary task for fine segmentation of lesions.

It is different from the deep learning method that uses two networks to perform tumor segmentation and classification in a two-stage framework. Yang *et al*. [306] used a multi-task method to design a deep learning model that could simultaneously segment and classify skin lesions. Among them, a two-stage classification model was used to improve the performance of classification and the model verified the feasibility of multi-task learning on a public dataset. Li *et al*. [307] proposed a deep learning framework that consisted of two FCRNs to simultaneously output segmentation results and rough classification results. Song *et al*. [308] designed an end-to-end approach to design a three-stage multi-task structure, and simultaneously performed lesion boundary detection, lesion segmentation, and disease classification tasks, which was superior to the state-of-the-art level under an improved loss function. Designed to improve classification and segmentation performance at the same time, Jin *et al*. [309] also proposed a similar three-stage network to aggregate cascading knowledge and transfer learning knowledge for different tasks. This method could avoid weight experience selection for different learning tasks. Wang and his team [310] designed a multi-task model for the diagnosis task of melanoma, embedding the lesion structure information in the skin lesion segmentation task into the lesion recognition task, and using the lesion type information in the recognition task to assist the pixel-level segmentation performance. The reliability of skin lesion analysis can be improved by integrating clinical knowledge into a deep learning architecture. The resulting automated system presents a three-level cascade model structure to further improve the representation of a multi-task melanoma diagnostic model. The task of segmentation and disease classification is adopted by most multi-task models. However, for the task of segmentation of lesions, the label of disease identification brings more labeling information to the multi-task model, which may not conform to the actual medical diagnosis logic. Liu *et al*. [311] used the detection of lesion edges as an auxiliary task of lesion segmentation, which could guide the segmentation model to pay more attention to the external edges of lesions and achieve better performance on the open data set.

As we know, CNN model constantly reduces the dimension of complex nonlinear features through convolution operation to capture deep linear features of data, which is one of the most efficient methods to find effective features of data. However, this abstract convolution operation comes at the expense of model interpretability. In the field of medical diagnosis, which is important to human health, professional doctors will never acknowledge the performance of opaque deep learning models. In the current field of machine learning, feature checking is frequently used to visualize model results during model reasoning. Popular strategies are the ability of class activation maps to highlight the image regions that contribute most to the feature maps at the model output level, or even to obtain the attention-discriminating deep features at the model feature extraction level. However, these methods can only be used after the training of the model and have no effect on changing the existing training model in time. In this case, some multi-task models [312] try to improve the interpretability and performance of CNN at the same time, which can guide the model with correct discriminant features in the training stage of the model. Barata *et al*. [127] imitated the diagnosis procedure of dermatologists and took the feature map extracted from CNN image data as the coding task. The decoding task was to make a multi-stage diagnosis according to the fixed levels of skin lesions by using LSTM model. The input of decoding task is filtered by the attentional module, which is helpful to strengthen the enhancement effect of refinement category on model performance. The Coppola team [313] is doing something similar. The rule-based method of the seven-point checklist has been widely used by doctors in practice and has a certain reference value. Different from previous studies [314] that only mapped the output to the score of the checklist to get the final diagnostic result, they added an additional 7-point evaluation index diagnostic task branch in the checklist in addition to the original diagnostic task of diseases, so that the multi-task model constructed in this way can make the deep opaque model have some interpretability. At the same time, the model introduces a gated structure to automatically learn beneficial features, which can also be used to directly explain the information shared between different tasks.

The commonly used medical classification tasks and segmentation tasks are not uniform in the difficulty of specific research objects. The team hoped to supplement the more challenging segmentation task with the classification task, which is easier to achieve in the research objective. Such joint training method has been proved in the previous literature to achieve better cooperation between tasks. Kong *et al*. [315] extracted the deep feature graph on the same benchmark backbone network, and then proposed a cross fusion module of key elements for the fusion interaction of multi-task branches. The idea of a crossover is to mitigate the performance degradation caused by feature mismatches between classification and segmentation tasks, since it is significantly easier to obtain the correct deep features for classification tasks. At the same time, in terms of loss setting, additional cross-supervision was also set for the branches of the classification task, in order to make the classification task with excellent performance as the dominant branch of the multi-task model. In the area of skin diseases, multi-task models are likely to require more tags than single-task models, which will undoubtedly increase the workload of professional physicians. Chu *et al*. [316] proposed a weak label task as a branch of multi-task, which can achieve 5% accuracy improvement and model acceleration effect compared to the original single task. Specifically, they use the existing segmentation labeling and rely on k-means clustering algorithm to generate weak labels for lesion classification, so as to optimize the two-loss functions. A variety of task combinations is an important factor affecting the performance improvement of multi-task models. In the field of skin diseases, focus segmentation tasks and disease recognition tasks are widely used. Partly because these two tasks are really necessary in clinical practice. They can greatly assist professional doctors in the diagnosis of skin diseases, save costs and reduce the subjectivity of doctors' diagnosis. On the other hand, the result of segmentation of lesions can provide a region of interest, which can add a layer of attention supervision to the task of disease identification. Similarly, the image-level label's output by the disease recognition task can guide the pixel-level classification of the lesion segmentation task. In addition, Jin's team [317] designed the multi-task model based on the prognostic prediction of disease and the task of lesion segmentation. Accurate prediction of the treatment response of individual patients is crucial for personalized medicine, and the results of segmentation of lesions can enable prognostic task models to focus on the edges of lesions. Therefore, in their multitask model, two twin networks were constructed according to the input multimodal images before and after treatment. The prognostic task is composed of multi-scale features of subnetworks. On the basis of the mainstream focus segmentation task and disease recognition task, Song's team [307] constructed three multi-task branching models in order to explore the contribution of focus edge features in focus detection task to other clinical diagnosis tasks. A feature pyramid and a regional recommendation network are used to generate rough regional branch inputs. On open data sets, they demonstrate that the three-branch multitask network can indeed provide an overall performance improvement for segmentation and recognition tasks.

7. CHANLLENGES

Above, we have described in detail the development process of the image analysis field in skin disease diagnosis. Deep learning technology has attracted widespread attention in the field of skin disease diagnosis and has made obvious progress. On the internationally public datasets, researchers continue to use improved models and algorithms to set new highs in indicators. In some literature, the experimental results can almost reach the level of dermatologists. It goes without saying that this is a constantly evolving and challenging field. This chapter aims to summarize the phased progress and existing shortcomings of the current research through the collation of the above literature and the elaboration of the research content, and guide the focus of future research work. Deep learning is essentially a data-driven model. Judging from the current state of the literature, artificial intelligence faces two main challenges in the field of skin disease image analysis diagnosis: limitations of data and models.

Each research will use one or more datasets to verify the performance of the system, and the experimental results are usually a breakthrough improvement over the performance of existing methods. However, we should treat the results in the literature with caution, mainly because the performance of the CAD model is very dependent on the quality of the dataset. Without exception, the excellent results of the model system are based on the specific dataset specified by the experiment, which does not have strict applicability in the true sense. The performance of deep learning algorithms mainly depends on the quality of the image dataset. Just slight image disturbances are enough to affect the diagnosis performance of CNN [318]. This is common in different publicly available skin lesion datasets. This paper has conducted some investigations and sorted out the relevant dataset in the latest research literature. Challenges involved in the skin disease image analysis task from the aspects of the dataset and model structure can be found in Fig. (**9**). We found that the problems exposed by the dataset are as follows.

7.1. Limited Overall Size of Dataset

In the field of skin disease image analysis diagnosis, the largest publicly available dataset, MoleMap, has only 102,451 images. Compared with millions of deep learning datasets in other fields, the total amount of labeled medical image data is very limited. Therefore, the problem frequently exposed in medical tasks is the over-fitting of models caused by insufficient sample data. Since medical data is

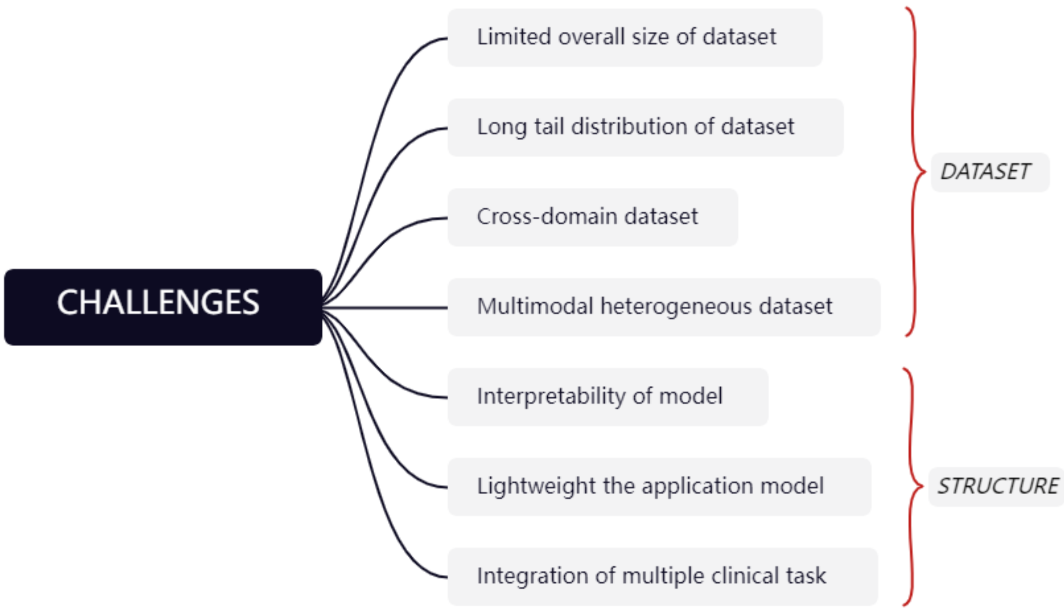

Fig. (9). Challenges involved in the skin disease image analysis task from the aspects of the dataset and model structure. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

more derived from patients' clinical visits, a considerable proportion of the data will not be used for research due to privacy agreements. At the same time, data labeling is also a costly task for dermatologists. Although there are currently some semi-automatic labeling methods, such as LabelImg [319] and LabelMe [320], they can effectively improve the efficiency of physicians' labeling. However, on the one hand, the proportion of manual participation in these methods is still too large, and the quality of the final data labels obtained is still far inferior to the labels that are completely manually labeled. At present, some researchers have designed some algorithms to realize the automatic acquisition of labels, but these methods of labeling have great limitations on data types [321]. The future work is to develop more robust algorithms to obtain data annotations. Similarly, transfer learning can also be considered as a branch of small-sample learning, in which pre-trained models are proven to be feasible for different tasks. The specific combination of models mainly depends on the actual data set size of the specific task and the similarity between different data sets. For example, Abubakar *et al*. [327] proposed that SVM works best as a classifier, but Hosny *et al*. [328] verified on more than one dataset that the accuracy of the SVM classifier was not high compared to other classifiers. In addition, another method of transfer learning is to freeze part of the deep network and train the remaining parameter layers. Although this method can alleviate the dilemma of the data set to a certain extent, it actually uses additional data information, not a small sample in the true sense, and deep features will still be subject to training limitations. We believe that in the two-stage implementation, the latter measure can only be regarded as an additional improvement in terms of performance. The real future research focus should be on how to achieve a better representation of small sample data.

In contrast, it was much easier to obtain large-scale unlabeled skin data. Some unsupervised CAD models [322, 323, 324, 325] have been proposed to complete special diagnostic tasks and can also achieve good diagnostic results. Unsupervised learning can reduce the dependence on large-scale labeled data, while achieving good performance. As far as the current research results are concerned, the performance of unsupervised learning on the same task is much lower than that of supervised learning on the same level of the data set. However, it is mentioned in the existing literature that the method of unsupervised learning can be applied to more practical tasks, for example, as a preprocessing step of supervised learning. At present, small sample learning methods have been applied to the field of auxiliary diagnosis of skin diseases [326]. Among them, metric learning is the most common. This method divides the task into two stages, representation and measure.

In addition to the unsupervised methods mentioned above, self-supervised learning can also provide another way to solve this kind of problem. Self-supervised learning models usually need to design appropriate auxiliary tasks from the data level according to the actual target tasks. This auxiliary task is designed to find a valid feature representation that allows large quantities of unlabeled data to be obtained. In this way, we can use a large amount of data to feed our model, so that our model can achieve generalization performance. Specifically, the effective features extracted by the auxiliary task of construction can be used as the pre-trained weights of our target model, which is the most popular operation at present. To some extent, this approach does not necessarily lead to a steady improvement in the actual model performance. The reason is that the features learned by the subjectively constructed auxiliary task do not match the needs of the target task. Such pre-trained weight will only lead to the misleading direction of the model at the beginning of the training stage, thus missing the needs of the features. In order to meet the needs of practical tasks in different fields, the existing auxiliary tasks for self-supervised learning are generally divided into three types, among which context-based auxiliary tasks are suitable for semantic domain, while time-sequential auxiliary tasks have relatively high data requirements for research. The auxiliary task based on contrast is used to extract features that measure the similarity between samples, that is, the recognition ability of sample features. This kind of auxiliary task has very loose requirements on the data and model of the target task, so it has attracted the attention of many scholars. After the investigation, self-supervised learning has not yet been applied in the diagnosis of skin diseases, which means both opportunities and challenges. We believe that it will not be long before suitable self-supervised auxiliary tasks are designed to improve the performance of the model.

While some tasks can be accomplished by cleverly designing unsupervised tasks, supervised tasks are more accurate and robust in terms of performance. Therefore, in more studies, we will choose the data-enhanced way to feed the model. Data enhancement, in a sense, uses a known label to generate different input data, which can increase the training data of the model, equivalent to expanding the original data set. The details of data enhancement are covered in Section 3 of this paper. However, relying too much on data enhancement will undoubtedly sacrifice part of data quality and make the model produce a certain degree of overfitting to harm the generalization performance. In actual research tasks, scholars need to reasonably induce the model to train more reasonably enhanced data according to the characteristics of data sets and research objectives of their own tasks, so that the limitation of too small data sets can undoubtedly be alleviated. At present, there is also a part of the literature that describes an algorithm to automatically find the best combination of data enhancement methods. Similarly, all possible parameter combinations are summarized into a hyperparameter space, and then different search algorithms are used to find the best combination of performance indicators. The actual effect of this algorithm often fails to meet the expectation of the algorithm because it cannot afford the expensive calculation cost. We believe that a more reasonable processing method is to consider the data level processing based on the research objective, screen out the data enhancement method that meets the expectation, and then search for the local optimal in the limited space with the help of the automatic parameter seeking algorithm, which is promising at present.

7.2. Long Tail Distribution of Dataset

A common problem in dermatological diagnosis tasks is the long tail distribution of the dataset. The reason is that, on the one hand, there are natural differences in the incidence of different skin diseases, and on the other hand, there are artificial differences in the later collection of dermatology departments of different diseases. Benign skin lesions generally have more cases than malignant lesions. Basal cell carcinoma, squamous cell carcinoma and melanoma together account for about 98% of all skin cancers. If this kind of problem is serious, the CAD model will be mistakenly introduced into the prejudice of the data category [329], that is, more objects are judged as the category of a large sample. Unbalanced datasets differ in their measures. For example, the imbalance of accuracy is defined according to the proportion of data between categories, which can better describe the performance of the model. To solve this problem, the current literature often balances the overall training set samples from the data level and the algorithm level. The data level is generally simple and rude, and the training set is balanced based on the idea of data sampling and data synthesis. GAN can synthesize images consistent with the distribution of the target dataset, so it has attracted a lot of attention in the field of computer vision in recent years. These synthesized images can be used as additional data of the CAD model to alleviate the quality of the dataset. In this way, many datasets based on small samples and unbalanced can also show good performance in the CAD model [330]. However, in the context of skin disease diagnosis, the synthesized image is different from the original image after all, and lacks reasonable interpretability. It is necessary to avoid the idea of mixing the synthesized image directly into the actual image set. However, the data-level method easily leads to overfitting of the training model, because the synthesized data is still distributed in the area enclosed by the known small sample data set. However, in the actual feature space, the true distribution of small sample data may not be limited to this area. Performing appropriate interpolation operations outside the bounding range may be able to achieve better data enhancement effects. From the perspective of algorithm, there are literatures that introduce an additional sampling rate in the input sampling stage of model training, so that relatively balanced data sets can be obtained to alleviate the category bias of the model no matter whether it is a small sample over-sampling or a variety of original under-sampling. However, the specific sampling strategy depends on the composition of the original data set, because repeated sampling for the same sample and sparse sampling for large samples will reduce the performance of the model. Scholars have found that starting from the algorithm level to increase the model's misjudgment and penalty loss method for small samples, it can eliminate the model's bias on the small sample category [143]. This method essentially produces a new data set distribution, but it can avoid the reuse or waste of the training set to a certain extent. For example, the two artificial parameters in Focal Loss correspond to two small branches of this type of method. One branch is to focus on the number of data to weight the loss during training. Specifically, the more categories, the less weight the corresponding input tagged with that category will lose during training. This weight value is inversely correlated with the quantity ratio in the literature, but it needs to be adjusted in different specific tasks. The other branch is more like a small skill, which is to increase the weight of loss of difficult samples in the process of model training, so that the model can focus more on the feature recognition of difficult samples and learn accurate features that are more in line with the research objectives. However, the design of a balanced penalty matrix is a complex task, and how to set an applicable penalty matrix for different task types is one of the future works. In addition to the above-mentioned common methods, there are still some other solutions. The integrated method achieves rigorous discrimination by comparing the results of the original training set and the "balanced" training set. In the unbalanced binary classification task, from a new perspective, try to design the task as anomaly detection of small sample categories. Anomaly detection is usually used to do unsupervised tasks, and its research object is usually a large number of single categories of normal data, with little or no abnormal data. Therefore, the anomaly detection task can meet the actual task of extreme data set structure. At present, traditional machine learning and deep learning methods exist in the literature, among which traditional machine learning mainly relies on statistics, distance, density and other indicators. The deep learning model represented by autoencoder is the mainstream method to solve this kind of problem. The input and output of the autoencoder are composed of the same data, and the low-dimensional features are obtained by dimensionality reduction of the convolution block. Then, the similarity between the feature restored image and the original input is used to evaluate the effect of the anomaly detection model. However, at present, the anomaly detection task is only applied to the task of extreme data sets, and the actual effect is not much better than supervised learning. We believe that the combination of this anomaly detection method and the integration idea will have good scalability in future work.

7.3. Cross-Domain Dataset

The sources of public datasets used in skin disease diagnosis studies are rather confusing, this brings serious cross-domain problems to the diagnostic model [20]. Due to the skin color or skin quality of different races, the representation of the same skin disease may have a large gap, so the acquired image data will inevitably bring challenges to the generalization of the model [161]. At the same time, the resolution of images acquired by different institutions lacks uniformity. An extreme resolution will hinder the feature learning process of the model, and will cause the model to not maintain the same performance on datasets of different resolutions. Therefore, it is very necessary to preprocess the image of the dataset. However, it is concluded in actual research that only considering the preprocessing operation is not enough at all. The current literature summarizes this type of problem as domain adaptation, which is mainly reflected in: the classification boundary of the classifier directly trained on the source domain cannot distinguish the samples of the target domain well. The actual methods presented in the current literature can be roughly divided into three types, which are the migration from the source domain to the target domain at the sample, feature and model levels, respectively. The method of sample migration needs to find similar data in the source domain, adjust the weight of the similar data by relying on manual experience, and finally retrain the classifier on the resampling sample set. Feature migration simply means replacing the migrated object with the common features of the sample. The corresponding idea is mainly an adaptive method based on features [104]. Map the features of different domains to the same space and achieve the smallest degree of discrimination. This processing method can also be adapted to unsupervised target domain learning. Model transfer is the mainstream of research at the present stage. The specific approach is to further learn the model through a small amount of target domain data on the basis of source domain training, so as to obtain good model performance. In addition to considering knowledge transfer across domains, there are research design models to learn domain-invariant features for detection tasks. Domain invariant features have strong sample inclusion and are general features between categories. The model includes adversarial learning module and a consistent regular module. The adversarial learning module is designed to select the adversarial classifier with the maximum and minimum cross-domain classification error from a set of classifiers, and then train the maximum difference between them. Finally, the consistent regular module monitors the minimal geometric distance between the pixel - and instance-level adversarial module outputs. This method can be applied to detection tasks with complex backgrounds and uneven imaging. In our opinion, finding an appropriate learning strategy to minimize the difference in the distribution of features in different domains is one of the mainstream directions for future work in this area. In addition, there are only a few commonly used skin diseases in the existing research. In terms of actual skin disease diagnosis tasks, such research results may not be generalized. Future research needs to be more implemented in other types that are in demand. A highly diverse skin disease dataset is of great significance for the construction of an effective CAD system.

7.4. Multimodal Heterogeneous Dataset

The visual features of the diseased area of the skin are the most critical diagnostic factor in the diagnosis process of skin diseases. Most of the existing researches is based on public medical image datasets, among which dermoscopy images with a single modality account for the majority. The classification criteria of skin diseases are very complicated, which makes visual features have great limitations in the diagnosis of many similar diseases. Therefore, some current researches have begun to shift in the direction of multi-source heterogeneous numbers. Different image data types often contain feature information of different dimensions [24]. Clinical images can characterize macroscopic visual characteristics: the original distribution and size of skin lesions, dermoscopy images provide a standardized view to characterize some colors and textures more clearly, and histopathological images focus on the internal structure of the cell level. At present, the research on various image data types usually obtains the types of diseases from the visual end-to-end based on the image input, and all have reached the benchmark level, but such research is not comprehensive enough. In addition, there is some information that cannot be represented visually. Other doctor-patient information (such as medical history, social habits, clinical metadata, etc.) is clinically important in the diagnosis of skin cancer. This information can provide clinicians with robust clinical guidance beyond the imaging features used by deep learning algorithms in actual examinations [331]. Therefore, if the deep learning algorithm used for the diagnosis of skin cancer is only based on medical images and ignores the key role of the patient's corresponding clinical information, there are problems to a certain extent. A previous study [332] demonstrated that the performance of both beginner and skilled dermatologists improved with the availability of clinical information, and their performance was better than deep learning algorithms. It can be imagined that the multi-modal heterogeneous model structure based on heterogeneous datasets can be favored by experts. One of the first prerequisites of multimodal models is data registration. Different modalities need to be registered according to the individual or even the part of the patient, which further increases the difficulty of data collection. In research, text data is often processed by natural language processing and data mining methods [333]. A data fusion algorithm needs to be developed to provide the final prediction of a skin cancer diagnosis. The algorithm can combine features composed of clinical information with imaging features from deep learning models. Unfortunately, in most publicly available skin lesion datasets, patient history and clinical metadata are missing. Even if some studies use patient metadata, the overall accuracy of the algorithm has been improved to a certain extent. However, these data only contain basic information such as the patient's age, gender, and diseased location, and only have certain guiding significance for a small number of disease types. Other textual medical history information with high general adaptability (such as: duration of illness and the algorithm for the assisted diagnosis of skin diseases has not yet been discovered.

In the existing literature on multimodal fusion, it is often used to define the level of model fusion, which usually includes early fusion, late fusion, and hybrid fusion. The early fusion can combine the original features of various modal images together, but compared with the deep features, the original features have a lot of redundant information and are mainly global features, which does not significantly improve the feature recognition of difficult samples. This is done by superimposing the channels of input data. This model is simple to implement and generally consists of only one complete branch, which is the most common form of fusion. Different from the early fusion, the late fusion will obtain the independent deep features of different modes, and more emphasis is placed on the control of total loss to realize the prediction decision optimization, so as to avoid the error accumulation in a single classifier [334]. Such models usually have more than one full branch, and the number of parameters in the model is multiplied. In terms of performance, it is easy to lose the advantages of multi-modal data sets. Proper improvement of post-fusion strategy will have good performance in specific tasks. As to medium-term fusion, can choose the location of the fusion, for the specific research tasks can bring very big flexibility, it is the image into a high dimensional feature vector expression, and in the middle layer model in specific ways, such as training model is easy to get to the commonness between different modal, the so-called common, it is to identify specific research characteristics. According to our survey, current studies on multimodal models are focused on exploring strategies for more ingenious intermediate fusion. For example, in the latest medical task literature [346], the basic experimental effects of different fusion forms were first explored, and then an intermediate branch was added to realize the fusion of different stages on the basis of the intermediate fusion, which was proved to provide good results in the final experiment. Although this paper provides a new attempt for mid-term fusion research in the field of medical diagnosis, it also exposes the shortcomings of artificially adjusting the weight parameters of different stages. However, scholars have not been able to reach an agreement on which layer of the model can bring the best results. Existing methods often use methods such as concatenation and element-wise products to map multi-modal information features to the same multi-modal model space dimension. But in other fields of research, such as semantic analysis, multimodal fusion has reached a more mature stage. At that time, we learned from related algorithms (such as based on matrix [335, 336, 337], based on attention [338, 339, 340], fusion matrix and features [341, 342, 343, 344], other methods [345]) can bring better results in the diagnosis of skin diseases to some extent. Many subsequent discussions on the optimal solution of model fusion and modal combination methods [347] have also become the direction of research efforts.

7.5. Interpretability of Model

In terms of model development, most of the current research trends focus on relying on CNN models to extract deep features between data. Although the deep learning method represented by CNN has better diagnostic performance in most tasks [348], people may not understand how CNNs with internal opacity determines the output, that is, such deep features are in a strict sense. Lack of scientific explanation [111].

At present, the mechanism of output results of most deep learning models is still a puzzle for researchers engaged in artificial intelligence. Therefore, especially in the high-stakes field of health care, it is almost impossible for professional physicians to apply such models to aid in diagnosis. In order to effectively help clinicians, deep learning algorithms need to provide semantic explanations for skin damage predictions, not just confidence scores. Also, it is different from the interpretable models recognized by academia, such as decision trees, linear models, etc. Existing interpretable AI algorithms can be divided into two categories: ad-hoc and post-hoc. Ad-hoc interpretation, that is, prior to model training, researchers use known prior interpretability characteristics to design a model with interpretability. Such models are interpretable in nature, whether from feature extraction or decision diagnosis. At the same time, the model can be tuned or guided to focus on areas or features that are actually interpretable during the training phase. This is very advantageous for the training of complex models [348]. Unlike previous studies that sought to guide deep learning models through some prior clinical information structures, Pintelas' team [349] proposed a more thorough and fully interpretive model. The fully interpreted model is designed to provide transparent performance decisions for the average person, not just explanations based on professional background. This fully interpreted model can extract the local texture of the input image, relying only on the statistical concepts of mean and variance, which can be easily understood by ordinary people. This method can construct a set of general feature structure systems by weight ratio, and combine linear machine learning algorithms to obtain a diagnosis performance not weaker than deep learning model.

Post-hoc interpretation is to explain the deep learning model in the model reasoning stage. In contrast, the model training and model interpretation are two independent stages, and the interpretable structure or feature introduced will not affect the training performance of the model [350]. At present, the mainstream research in this direction is the perturbation method and gradient method. The perturbation method is to measure the influence of model precision decline through local displacement, fuzzy and other operations. The gradient method is mainly based on the convolution layer thermal map. Simple proxy models have a more transparent interpretation. In general, nonlinear and complex models are replaced by local interpretable models. A local proxy model can be used to observe the features of the model well. Stieler *et al*. [351] used three different proxy models to observe the explanatory features of the ABCD criterion, and evaluated the features that have the greatest influence on the deep learning model. But such ideas must be used with care because they lack basic facts and methods. In particular, many studies have proposed methods that are primarily influenced by visual features, such as high contrast and reduced noise. Many of these denoised salient graphs may result in strong biases that are inconsistent with the true interpretability of the underlying deep learning model. CAM thermogram is based on gradient method to calculate the class fraction of convolution layer, that is, feature graph. Therefore, this method can quantify the abstract features learned from the middle layer, and can obtain the visual result of the combination of feature graph, saliency graph and input dimension. Jiang *et al*. [352] set up the CAM module before the output level to provide the attention distribution of the diagnostic model of visual histopathological images. Olah *et al*. [353] proposed to perform gradient ascent on a unit in the middle level of the model to visualize the sensitivity of a single unit to global performance. Clearly, this model interpretation method can show significant regions in a given sample, but may lack specificity for machines and people.

In conclusion, most of these methods can only provide approximate explanations, and cannot fully reflect the true behavior of the model, which is contrary to the rigorous requirements of the field of medical assisted diagnosis. How to establish the interpretability of deep model diagnosis has also become one of the obstacles to the advancement of research applications. The fusion of handcrafted features or multi-level features in the model may provide some reference interpretability for deep learning models. For example, Kawahara *et al*. [156] provide semantic interpretation of network predictions according to ABCD standards or 7-point checklist. Therefore, future research work may start from studying the etiology and visual characteristics of skin diseases, and then designing a deep network with domain knowledge for specific tasks. In this way, better performance can be expected.

7.6. Lightweight the Application Model

When our scholars design advanced algorithms or models, they will undoubtedly target various evaluation indicators on a certain public data set, such as accuracy, recall, specificity in classification tasks, Dice, Jac in segmentation tasks, mAP in detection tasks, etc. Under the background of blindly pursuing model "index" performance, the existing models have more and more parameters, more computation, and complex structure, but the running speed keeps decreasing. In reality, model algorithms with high "indicators" are only allowed to re-achieve the indicators in the researcher's working environment. Especially for those model algorithms that need practical applications, such as skin disease diagnosis, this is a huge problem for non-professional hardware devices.

For example, the application of a skin disease auxiliary diagnosis system is generally deployed on image data acquisition equipment. While devices such as dermoscopy images rely on mirrors with physical magnification, clinical images are more mobile cameras or mobile phones. It is possible to add external processor hardware, but the original mobile devices obviously do not have the memory and reasoning power required by the model algorithm. Therefore, it is very urgent to design a model that cannot lose the high "index" performance of the original model, but also has the characteristics of lightweight and landing. This is an important and very active field. At present, the general operation of model lightweight mainly includes several aspects, such as hardware, platform and algorithm. This paper only realizes the lightweight of the model at the algorithm level.

In recent years, many scholars have made many achievements in deep learning compression and acceleration, and a variety of neural network lightweight algorithms have emerged. One of the outstanding significances is the optimization of network structure. It was first proposed in the InceptionV2 model that a large convolution kernel was replaced by two small convolution kernels in series. Without affecting the effect of the model too much, the number of parameters can be reduced to about half of the original. This discovery can be embedded into most network models without loss and is the first work of model lightweight research. Since the series of convolution kernels can be improved, can parallel connections also be considered? Subsequently, in the later version of InceptionV3, The Google team proposed to replace a normal convolution kernel with two parallel asymmetric convolution kernels. InceptionV3 splits a 7x7 convolution into a 1x7 convolution and a 7x1 convolution. Under similar convolution effects, the number of parameters is greatly reduced, and the diversity of convolution is also improved. Later scholars came to the conclusion that the same is true for n x n convolution kernels. Instead of simply stacking the network layers, Inception series networks reduce the number of parameters by changing the size of the convolutional kernel. After investigation of the literature in recent years, it is concluded that the improvement trend of convolution kernel size is to use small convolution kernels to replace large convolution kernels. Although the combination of small convolution kernels can indeed reduce the number of parameters in the model, it will also bring the problem of too small feature area capture. In the deep learning model, the later network hierarchy will capture deeper abstract features. Therefore, in network design, lightweight modules based on small convolutional kernels are usually placed in shallow structures, while large convolutional kernels are still used to increase the receptive field in deep structures. Inspired by this, more construction methods of the network layer are designed to realize the compression and acceleration of the model. ShuffleNet and MobileNet are outstanding achievements in the hierarchical lightweight of network structures. Firstly, the model presents the concept of grouping convolution, which is used to group the channels of the feature map into N groups. Each filter can operate in each group. However, the grouped output channels will lose the opportunity of information interaction. The solution in ShuffleNet is shuffle Operation to strengthen the connection between channels. On this basis, we directly consider the extreme case of grouping convolution, where a grouping consists of only one convolution. In the second stage, 1×1 convolution kernel is used to transform the output channel, and the information fusion between channels is also carried out. This method is called depth separable convolution. In the field of skin disease diagnosis, depthwise separable convolution has been widely used in various task networks, which has quite good applicability. Directly reducing the number of convolutional kernels can reduce the number of model parameters, but it will lead to the reduction of the output feature graph and the feature expression ability of the network. A similar solution is the idea of density. While reducing the number of filters at each layer, the feature map output at each layer is reused before the input of the network at each layer. Pooling operation can reduce the size of feature channels at will to reduce the cost of convolution multiplication. Compared with convolution operation, pooling layer structure has no parameters that need to be updated. At the output level of the network, some networks will replace a large number of parameters of the full connection layer with the global average pooling layer. Model pruning is an important research branch of model lightweight. The general idea is to set the unimportant parameters in the weight matrix of model update to 0. In order to maintain the original performance of the model, it is often necessary to iteratively prune the model in small steps on a high-performance processor to find the most suitable lightweight model.

In addition to modifying network hierarchy, model distillation is a method to improve model compression and acceleration in model algorithm training, and it is also a very common way before the model algorithm is applied. The essence of distillation is to fit two models of varying degrees of complexity so that a simple student network can learn from a complex and high-performance teacher network knowledge capable of handling research tasks. The general specific approach is to define the soft label task, fit the loss of the student network and the teacher network in this task, and then make predictions on the target research task. The current literature basically studies this kind of task from two perspectives, one is the selection of teacher network, and, more importantly, the definition of soft label. According to the literature experience description, teacher network and student network can be completely different network structures, but generally similar network structure, distillation effect will be better [354]. The famous Sanh team combined three different soft tag losses to fit two similar networks. The distilled model is 60% faster and 60% smaller than the original model. More than 95% of performance is preserved on public tasks. With little performance loss, the model can be compressed and accelerated. [355] Among them, attention module level and feature hiding level have the most potential cooperation. After comparative experiments, they concluded that the fitting method corresponding to deep structure and shallow structure is easier to obtain high-performance student distillation model. In the structural design of the model, neural structure search (NAS) is also used to find the network structure that conforms to the compression and acceleration constraints, which will result in a high computational cost, which will not be described in more detail here. The future research direction is to select and design an appropriate teacher network and soft label combined with the research task, which plays a decisive role in the effect of model distillation.

7.7. Integration of Multiple Clinical Task

It can be seen from more and more recent studies that there is a certain correlation between medical diagnosis tasks of different application backgrounds. Utilizing the correlation between the segmentation task and the classification task to construct an integrated framework has become a hot topic in the diagnosis of skin diseases [273]. The core problems faced by classification tasks are the similarities within the class and the confusion of the background. There is a lot of artificial interference around or even inside the focal area of skin diseases. These uncertain and complex backgrounds are likely to reduce the model's ability to extract similar information from the same skin lesion category. The problem faced by the segmentation task is more that the segmented background and foreground generated by the model cannot be perfectly integrated at the edge, which is essentially a problem of poor pixel classification. Whether the feature information of skin lesion edge can be extracted is an important influence on the effect of the lesion segmentation model. However, due to the uncertainty of skin edge, including shape, color, texture, etc., it is a great challenge to deal with edge pixels in the skin segmentation task. The current mainstream processing method in the literature is to use the output of the previous network to modify the latter network to obtain more accurate results. Since the ultimate goal of medical tasks is to make accurate diagnosis possible, some studies have put forward the advantages of combining classification learning and comparative learning model, which can be summarized as expanding the distance between classes and reducing the distance within classes. Distinguishing different categories focus on finding different features between different categories. However, in practice, data belonging to the same category may not have exactly the same features. For example, different expressions on the same person's face will learn different features. However, in contrast, learning, similar features of the same category are extracted, but such features cannot be applied to the classification. Therefore, the two types of tasks complement each other to some extent. The features they focus on may overlap, but there is always a bias. This complementarity ensures that the combination of the two types of tasks theoretically yields an overall performance improvement. In other words, considering the actual model integration design, the adaptability between different modules in the model and the difference of training environment will bring no small challenge to the integration of different tasks. At present, the design of such an integration model still depends on the experience of researchers engaged in related projects for many years, and some automatic search algorithms still cannot be deployed in various fields based on a limited hardware environment. There are also attempts to perform different tasks in a multi-task model at the same time, and better results can be obtained by effectively sharing part of the weight parameters of the model. Sharing weights will undoubtedly bring more serious hardware costs. Such a model will greatly increase the difficulty of model lightweight in practical application and limit its real research significance. Regardless of the extra cost, weight sharing will fail to achieve the desired effect due to the uncertain learning direction during model training, even worse than the effect of any single task. It is true that the above method of utilizing task relevance can greatly improve the effectiveness of the task, but much of this relevance is reflected in the combination of the model accompanied by: 1) the construction of the same magnitude loss function; 2) the learning speed of the task balance: 3) the pertinence of target tasks (tend). These uncertainties in the process of model training will have different effects on the multi-task model, or even completely different. How to efficiently utilize this correlation will be a major challenge for researchers. These uncertainties in the process of model training will have different effects on the multi-task model, or even completely different. In practice, the main advantage of multi-task learning lies in merging models and reducing the number of model parameters. As for the performance boost to the target task, theoretically, similar tasks can improve each other's performance, which may not be so obvious. This is because it is not easy to shrink models while maintaining the performance of each model. Adjusting the network structure to fit the multi-task model is cumbersome and depends heavily on processing experience. In the literature, the dynamic assignment of weight loss is explored by many people [356]. In terms of the research direction of multi-task model structure, other mature fields, such as remote sensing and voice, will bring a lot of reference value to the tasks in the field of medical image intelligent analysis. In the future, a more optimized general-template idea can be constructed based on the actual characteristics of the medical diagnosis field.

CONCLUSION

This paper describes the development of image analysis technology in computer-aided diagnosis. We briefly introduce skin disease image acquisition data and preprocessing technology. Starting from the actual needs of skin disease diagnosis tasks, the classification, detection and segmentation tasks of skin diseases are reviewed and expanded more methods in skin disease diagnosis tasks, especially deep learning. While reviewing the research results in recent years, we also discussed the problems currently exposed, and provided guidance for future follow-up directions. From this review, we can observe that many image analysis methods have been proposed in various fields of skin disease diagnosis, and equivalent or better diagnostic performance has been achieved on the experimental skin disease dataset. However, we should be aware that there is still a lot of room for improvement in aspects of the dataset and model structure, before image intelligent analysis technology is applied to actual clinical systems. We hope that this work will be a valuable guide for researchers to make progress in this field.

CONSENT FOR PUBLICATION

Not applicable.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

[1] Hay RJ, Johns NE, Williams HC, *et al.* The global burden of skin disease in 2010: An analysis of the prevalence and impact of skin conditions. J Invest Dermatol 2014; 134(6): 1527-34. http://dx.doi.org/10.1038/jid.2013.446 PMID: 24166134

[2] Han SS, Kim MS, Lim W, Park GH, Park I, Chang SE. Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm. J Invest Dermatol 2018; 138(7): 1529-38. http://dx.doi.org/10.1016/j.jid.2018.01.028 PMID: 29428356

[3] Wernli KJ, Henrikson NB, Morrison CC, Nguyen M, Pocobelli G, Blasi PR. Screening for skin cancer in adults: Updated evidence report and systematic review for the US preventive services task force. JAMA 2016; 316(4): 436-47. http://dx.doi.org/10.1001/jama.2016.5415 PMID: 27458949

[4] Lowe DG. Distinctive image features from scale-invariant keypoints. Int J Comput Vis 2004; 60(2): 91-110. http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94

[5] van de Weijer J, Schmid C, Verbeek J, Larlus D. Learning color names for real-world applications. IEEE Trans Image Process 2009; 18(7): 1512-23. http://dx.doi.org/10.1109/TIP.2009.2019809 PMID: 19482579

[6] Carli P, Quercioli E, Sestini S, *et al.* Pattern analysis, not simplified algorithms, is the most reliable method for teaching dermoscopy for melanoma diagnosis to residents in dermatology. Br J Dermatol 2003; 148(5): 981-4. http://dx.doi.org/10.1046/j.1365-2133.2003.05023.x PMID: 12786829

[7] Argenziano G, Fabbrocini G, Carli P, De Giorgi V, Sammarco E, Delfino M. Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions. Comparison of the ABCD rule of dermatoscopy and a new 7-point checklist based on pattern analysis. Arch Dermatol 1998; 134(12): 1563-70. http://dx.doi.org/10.1001/archderm.134.12.1563 PMID: 9875194

[8] Menzies SW, Bischof L, Talbot H, *et al.* The performance of SolarScan: An automated dermoscopy image analysis instrument for the diagnosis of primary melanoma. Arch Dermatol 2005; 141(11): 1388-96. http://dx.doi.org/10.1001/archderm.141.11.1388 PMID: 16301386

[9] Abbasi NR, Shaw HM, Rigel DS, *et al.* Early diagnosis of cutaneous melanoma: Revisiting the ABCD criteria. JAMA 2004; 292(22): 2771-6. http://dx.doi.org/10.1001/jama.292.22.2771 PMID: 15585738

[10] Henning JS, Dusza SW, Wang SQ, *et al.* The CASH (color, architecture, symmetry, and homogeneity) algorithm for dermoscopy. J Am Acad Dermatol 2007; 56(1): 45-52. http://dx.doi.org/10.1016/j.jaad.2006.09.003 PMID: 17190620

[11] Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995; 20(3): 273-97. http://dx.doi.org/10.1007/BF00994018

[12] Ho TK. Random decision forests Proceedings of 3rd international conference on document analysis and recognition. 278-82.

[13] Altman NS. An introduction to kernel and nearest-neighbor nonparametric regression. Am Stat 1992; 46(3): 175-85.

[14] Kamiński B, Jakubczyk M, Szufel P. A framework for sensitivity analysis of decision trees. Cent Eur J Oper Res 2018; 26(1): 135-59. http://dx.doi.org/10.1007/s10100-017-0479-6 PMID: 29375266

[15] Esteva A, Kuprel B, Novoa RA, *et al.* Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017; 542(7639): 115-8. http://dx.doi.org/10.1038/nature21056 PMID: 28117445

[16] Zhao S, Xie B, Li Y, *et al.* Smart identification of psoriasis by images using convolutional neural networks: A case study in China. J Eur Acad Dermatol Venereol 2020; 34(3): 518-24. http://dx.doi.org/10.1111/jdv.15965 PMID: 31541556

[17] Liao H, Luo J. A deep multi-task learning approach to skin lesion classification. arXiv 2018.

[18] Codella NCF, Gutman D, Celebi ME, *et al.* Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic). IEEE Int Symp Biomed Imag 2018; 2018: 168-72.

[19] Zhang J, Xie Y, Wu Q, *et al.* Skin lesion classification in dermoscopy images using synergic deep learning. Int Conf Med Image Comput Comput-Assist Interven 2018; 2018: 12-20.

[20] Xie B, He X, Zhao S, *et al.* XiangyaDerm: A Clinical Image Dataset of Asian Race for Skin Disease Aided Diagnosis//Large-Scale Annotation of Biomedical Data and Expert Label Synthesis and Hardware Aware Learning for Medical Imaging and Computer Assisted Intervention. Cham: Springer 2019; pp. 22-31.

[21] Pal A, Chaturvedi A, Garain U, *et al.* CapsDeMM: Capsule network for detection of munro's microabscess in skin biopsy images. Int Conf Med Imag Comput Comput-Assist Interven 2018; 2018: 389-97.

[22] Marghoob A, Braun R. An Atlas of Dermoscopy. USA: CRC Press 2012.

[23] Day GR, Barbour RH. Automated melanoma diagnosis: Where are we at? Skin Res Technol 2000; 6(1): 1-5. http://dx.doi.org/10.1034/j.1600-0846.2000.006001001.x PMID: 11428935

[24] Haenssle HA, Fink C, Schneiderbauer R, *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): 1836-42. http://dx.doi.org/10.1093/annonc/mdy166 PMID: 29846502

[25] Pellacani G, Seidenari S. Comparison between morphological parameters in pigmented skin lesion images acquired by means of epiluminescence surface microscopy and polarized-light videomicroscopy. Clin Dermatol 2002; 20(3): 222-7. http://dx.doi.org/10.1016/S0738-081X(02)00231-6 PMID: 12074856

[26] Kopf A W, Elbaum M, Provost N. The use of dermoscopy and digital imaging in the diagnosis of cutaneous malignant melanoma. Skin Res Technol 1997; 3(1): 1-7. http://dx.doi.org/10.1111/j.1600-0846.1997.tb00152.x

[27] Menzies SW. Automated epiluminescence microscopy: Human *vs.* machine in the diagnosis of melanoma. Arch Dermatol 1999; 135(12): 1538-40. http://dx.doi.org/10.1001/archderm.135.12.1538 PMID: 10606065

[28] Benvenuto-Andrade C, Dusza SW, Agero AL, *et al.* Differences between polarized light dermoscopy and immersion contact dermoscopy for the evaluation of skin lesions. Arch Dermatol 2007; 143(3): 329-38. http://dx.doi.org/10.1001/archderm.143.3.329 PMID: 17372097

[29] Binder M, Schwarz M, Winkler A, *et al.* Epiluminescence microscopy. A useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists. Arch Dermatol 1995; 131(3): 286-91. http://dx.doi.org/10.1001/archderm.1995.01690150050011 PMID: 7887657

[30] The international skin imaging collaboration (ISIC). 2020. Available from: https:// www.isic-archive.com/ (Accessed on: 2, 2020).

[31] Dermofit. A cognitive prosthesis to aid focal skin lesion diagnosis. Available from: https://homepages.inf.ed.ac.uk/rbf/DERMOFIT/

[32] ADDI Project. PH2. Available from: https://www.fc.up.pt/addi/

[33] Goyal M, Knackstedt T, Yan S, Hassanpour S. Artificial intelligence-based image classification methods for diagnosis of skin cancer: Challenges and opportunities. Comput Biol Med 2020; 127: 104065. http://dx.doi.org/10.1016/j.compbiomed.2020.104065 PMID: 33246265

[34] MED_NODE Database. Dermatology database used in MED-NODE. Available from: http://www.cs.rug.nl/~imaging/databases/melanoma_naevi/

[35] Derm101. Available from: http://www.derm101.com/

[36] SD-198. Recognition of clinical skin disease images. Available from: http://xiaopingwu.cn/assets/projects/sd-198/

[37] Dermnet.. Skin Disease Atlas. Available from: http://www.dermnet.com/

[38] Atlasderm.. Dermatology Atlas. Available from: http://www.atlasdermatologico.com.br/

[39] Danderm. Available from: http://www.danderm.dk/

[40] Derm IS. Available from: https://www.dermis.net/dermisroot/en/home/indexp.html/

[41] Asan. Available from: https://figshare.com/articles/Asan_and_Hallym_Dataset_Thumbnails_/5406136/

[42] Molemap. Available from: https://www.molemap.net.au/

[43] Rubin's pathology: Clinicopathologic foundations of medicine. Pennsylvania, USA: Lippincott Williams & Wilkins 2008.

[44] Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. Histopathological image analysis: A review. IEEE Rev Biomed Eng 2009; 2: 147-71. http://dx.doi.org/10.1109/RBME.2009.2034865 PMID: 20671804

[45] TCGA. Available from: https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga/

[46] Poynton C. Digital video and HD: Algorithms and Interfaces. Elsevier 2012.

[47] Pratt W. Spatial transform coding of color images. IEEE Trans Commun Technol 1971; 19(6): 980-92. http://dx.doi.org/10.1109/TCOM.1971.1090769

[48] Ahmad T, Farou Z. Supervised learning methods for skin segmentation based on pixel color classification. Cent-Eur J New Technol Res Educ Pract 2021. [Epub ahead of print].

[49] Barata C, Celebi ME, Marques JS. Improving dermoscopy image classification using color constancy. IEEE J Biomed Health Inform 2015; 19(3): 1146-52. PMID: 25073179

[50] Hua Ng J, Goyal M, Hewitt B, *et al.* The effect of color constancy algorithms on semantic segmentation of skin lesions. Med Imag 2019; 10953: 10953.

[51] Gómez DD, Butakoff C, Ersbøll BKÆ, Stoecker W. Independent histogram pursuit for segmentation of skin lesions. IEEE Trans Biomed Eng 2008; 55(1): 157-61. http://dx.doi.org/10.1109/TBME.2007.910651 PMID: 18232357

[52] Celebi ME, Iyatomi H, Schaefer G, Stoecker WV. Lesion border detection in dermoscopy images. Comput Med Imaging Graph 2009; 33(2): 148-53. http://dx.doi.org/10.1016/j.compmedimag.2008.11.002 PMID: 19121917

[53] Norton KA, Iyatomi H, Celebi ME, *et al.* Three-phase general border detection method for dermoscopy images using non-uniform illumination correction. Skin Res Technol 2012; 18(3): 290-300. http://dx.doi.org/10.1111/j.1600-0846.2011.00569.x PMID: 22092500

[54] Iyatomi H, Celebi ME, Schaefer G, Tanaka M. Automated color calibration method for dermoscopy images. Comput Med Imaging Graph 2011; 35(2): 89-98. http://dx.doi.org/10.1016/j.compmedimag.2010.08.003 PMID: 20933366

[55] Schaefer G, Rajab MI, Celebi ME, Iyatomi H. Colour and contrast enhancement for improved skin lesion segmentation. Comput Med Imaging Graph 2011; 35(2): 99-104. http://dx.doi.org/10.1016/j.compmedimag.2010.08.004 PMID: 21035303

[56] Melinscak M, Prentasic P, Loncaric S. Retinal vessel segmentation using deep neural networks. VISAPP 2015; (1): 577-82. http://dx.doi.org/10.5220/0005313005770582

[57] Bisla D, Choromanska A, Stein JA, *et al.* Skin lesion segmentation and classification with deep learning system. arXiv 2019; 2019: 1-6.

[58] Jafari MH, Karimi N, Nasr-Esfahani E, *et al.* Skin lesion segmentation in clinical images using deep learning. Int Conf Pattern Recogn (ICPR) 2016; 2016: 337-42.

[59] Vala HJ, Baxi A. A review on Otsu image segmentation algorithm. Int J Adv Res Comput Eng Technol 2013; 2(2): 387-9. [IJARCET].
[60] Huang ZK, Chau KW. A new image thresholding method based on Gaussian mixture model. Appl Math Comput 2008; 205(2): 899-907. http://dx.doi.org/10.1016/j.amc.2008.05.130
[61] Khan HA, Iskandar DNF, Al-Asad JF, *et al.* Classification of skin lesion with hair and artifacts removal using black-hat morphology and total variation. Int J Comput Digital Sys 2020; 10: 2-8.
[62] Zhao R, Ouyang W, Li H, *et al.* Saliency detection by multi-context deep learning. Proceedings of the IEEE conference on computer vision and pattern recognition. 7-12 June 2015; Boston, MA, USA.
[63] Pereira S, Pinto A, Alves V, *et al.* Deep convolutional neural networks for the segmentation of gliomas in multi-sequence MRI//BrainLes 2015. Cham: Springer 2015; pp. 131-43.
[64] Talavera-Martinez L, Bibiloni P, Gonzalez-Hidalgo M. Hair segmentation and removal in dermoscopy images using deep learning. IEEE Access 2020; 9: 2694-704. http://dx.doi.org/10.1109/ACCESS.2020.3047258
[65] Badrinarayanan V, Kendall A, Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans Pattern Anal Mach Intell 2017; 39(12): 2481-95. http://dx.doi.org/10.1109/TPAMI.2016.2644615 PMID: 28060704
[66] Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille AL. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE Trans Pattern Anal Mach Intell 2018; 40(4): 834-48. http://dx.doi.org/10.1109/TPAMI.2017.2699184 PMID: 28463186
[67] Yan Z, Zhan Y, Peng Z, *et al.* Multi-instance deep learning: Discover discriminative local anatomies for bodypart recognition. IEEE Trans Med Imaging 2016; 35(5): 1332-43. http://dx.doi.org/10.1109/TMI.2016.2524985 PMID: 26863652
[68] Miao S, Wang ZJ, Liao R. A CNN regression approach for real-time 2D/3D registration. IEEE Trans Med Imaging 2016; 35(5): 1352-63. http://dx.doi.org/10.1109/TMI.2016.2521800 PMID: 26829785
[69] Celebi ME, Iyatomi H, Schaefer G, Stoecker WV. Approximate lesion localization in dermoscopy images. Skin Res Technol 2009; 15(3): 314-22. http://dx.doi.org/10.1111/j.1600-0846.2009.00357.x PMID: 19624428
[70] Wang H, Chen X, Moss RH, *et al.* Watershed segmentation of dermoscopy images using a watershed technique. Skin Res Technol 2010; 16(3): 378-84. http://dx.doi.org/10.1111/j.1600-0846.2010.00445.x PMID: 20637008
[71] Wang H, Moss RH, Chen X, *et al.* Modified watershed technique and post-processing for segmentation of skin lesions in dermoscopy images. Comput Med Imaging Graph 2011; 35(2): 116-20. http://dx.doi.org/10.1016/j.compmedimag.2010.09.006 PMID: 20970307
[72] Abbas Q, Celebi ME, Garcia IF. A novel perceptually-oriented approach for skin tumor segmentation. Int J Innov Comput, Inf Control 2012; 8(3): 1837-48.
[73] Emre Celebi M, Alp Aslandogan Y, Stoecker WV, Iyatomi H, Oka H, Chen X. Unsupervised border detection in dermoscopy images. Skin Res Technol 2007; 13(4): 454-62. http://dx.doi.org/10.1111/j.1600-0846.2007.00251.x PMID: 17908199
[74] Celebi ME, Kingravi HA, Iyatomi H, *et al.* Border detection in dermoscopy images using statistical region merging. Skin Res Technol 2008; 14(3): 347-53. http://dx.doi.org/10.1111/j.1600-0846.2008.00301.x PMID: 19159382
[75] Ünver HM, Ayan E. Skin lesion segmentation in dermoscopy images with combination of YOLO and grabcut algorithm. Diagnostics (Basel) 2019; 9(3): 72. http://dx.doi.org/10.3390/diagnostics9030072 PMID: 31295856
[76] Zheng L, Zhao Y, Wang S, *et al.* Good practice in CNN feature transfer. arXiv 2016; 2016: 1604.00133.
[77] Yu Z, Jiang X, Zhou F, *et al.* Melanoma recognition in dermoscopy images *via* aggregated deep convolutional features. IEEE Trans Biomed Eng 2019; 66(4): 1006-16. http://dx.doi.org/10.1109/TBME.2018.2866166 PMID: 30130171
[78] Rastgoo M, Garcia R, Morel O, Marzani F. Automatic differentiation of melanoma from dysplastic nevi. Comput Med Imaging Graph 2015; 43: 44-52. http://dx.doi.org/10.1016/j.compmedimag.2015.02.011 PMID: 25797605
[79] Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. J Big Data 2019; 6(1): 1-48. http://dx.doi.org/10.1186/s40537-019-0197-0
[80] Nyíri T, Kiss A. Style transfer for dermatological data augmentation. Proc SAI Intell Sys Conf 2020; 2020: 915-23.
[81] Chengchuang L, Chun S, Gansen Z, *et al.* Review of image data augmentation in computer vision. Comput Sci Appl 2021; 11(2): 13.
[82] Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic minority over-sampling technique. J Artif Intell Res 2002; 16(1): 321-57. http://dx.doi.org/10.1613/jair.953
[83] Zhang H, Cisse M, Dauphin YN, *et al.* mi xup: Beyond empirical risk minimization. arXiv 2017; 2017: 1710.09412.
[84] Inoue H. Data augmentation by pairing samples for images classification. arXiv 2018; 2018: 1801.02929.
[85] Yun S, Han D, Oh SJ, *et al.* Cutmix: Regularization strategy to train strong classifiers with localizable features. Proc IEEE/CVF Int Conf Comput Vision 2019; 2019: 6023-32.
[86] Shah V, Autee P, Sonawane P. Detection of melanoma from skin lesion images using deep learning techniques. Int Conf Data Sci Eng (ICDSE) 2020; 2020: 1-8.
[87] Perez F, Vasconcelos C, Avila S, *et al.* Data augmentation for skin lesion analysis//OR 20 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis. Cham: Springer 2018; pp. 303-11.
[88] Pham TC, Luong CM, Visani M, *et al.* Deep CNN and data augmentation for skin lesion classification. Asian Conf Intell Inform Database Sys 2018; 2018: 573-82.
[89] Al-Masni MA, Al-Antari MA, Choi MT, Han SM, Kim TS. Skin lesion segmentation in dermoscopy images *via* deep full resolution convolutional networks. Comput Methods Programs Biomed 2018; 162: 221-31. http://dx.doi.org/10.1016/j.cmpb.2018.05.027 PMID: 29903489
[90] Cubuk ED, Zoph B, Mane D, *et al.* Autoaugment: Learning augmentation policies from data. arXiv 2018; 2018: 1805.09501.
[91] Goodfellow IJ, Pouget-Abadie J, Mirza M, *et al.* Generative adversarial networks. Adv Neural Inf Process Syst 2014; 3: 2672-80.
[92] Cubuk ED, Zoph B, Shlens J, *et al.* Randaugment: Practical automated data augmentation with a reduced search space. Proc IEEE/CVF Conf Comput Vision Pattern Recogn Workshops 2019; 2019: 702-3.
[93] Li Y, Hu G, Wang Y, *et al.* Differentiable automatic data augmentation. Eur Conf Comput Vision 2020; 2020: 580-95.
[94] Shen S, Xu M, Zhang F, *et al.* Low-cost and high-performance data augmentation for deep-learning-based skin lesion classification arXiv 2021; 2021: 2101.02353.
[95] Goodfellow I, Pouget-Abadie J, Mirza M, *et al.* Generative adversarial nets. Adv Neural Inf Process Syst 2014; 2014: 27.
[96] Yi X, Walia E, Babyn P. Generative adversarial network in medical imaging: A review. Med Image Anal 2019; 58: 101552. http://dx.doi.org/10.1016/j.media.2019.101552 PMID: 31521965
[97] Wei J, Suriawinata A, Vaickus L, *et al.* Generative image translation for data augmentation in colorectal histopathology. Images//Machine Learn Health Workshop PMLR 2020; 2020: 10-24.
[98] Bissoto A, Perez F, Valle E, *et al.* Skin lesion synthesis with generative adversarial networks//OR 20 context-aware operating theaters, computer assisted robotic endoscopy, clinical image-based procedures, and skin image analysis. Cham: Springer 2018; pp. 294-302.
[99] Rashid H, Tanveer MA, Khan HA. Skin lesion classification using GAN based data augmentation. Ann Int Conf IEEE Eng Med Biol Soc (EMBC) 2019; 2019: 916-9.

[100] Bisla D, Choromanska A, Berman RS, *et al.* Towards automated melanoma detection with deep learning: Data purification and augmentation. Proc IEEE/CVF Conf Comput Vision Pattern Recogn Workshops arXiv 2019; 2019: 1902.06061.
[101] Pollastri F, Bolelli F, Paredes R, Grana C. Augmenting data with GANs to segment melanoma skin lesions. Multimedia Tools Appl 2020; 79(21): 15575-92. http://dx.doi.org/10.1007/s11042-019-7717-y
[102] Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset, a large collection of multi-source dermoscopy images of common pigmented skin lesions. Sci Data 2018; 5(1): 1-9. http://dx.doi.org/10.1038/sdata.2018.161 PMID: 30482902
[103] Goyal M, Hassanpour S, Yap MH. Region of interest detection in dermoscopy images for natural data-augmentation. arXiv 2018; 2018: 1807.10711.
[104] Ghorbani A, Natarajan V, Coz D, *et al.* DermGAN: Synthetic generation of clinical skin disease images with pathology. PMLR 2020; 2020: 155-70.
[105] Gu Y, Ge Z, Bonnington CP, Zhou J. Progressive transfer learning and adversarial domain adaptation for cross-domain skin disease classification. IEEE J Biomed Health Inform 2020; 24(5): 1379-93. http://dx.doi.org/10.1109/JBHI.2019.2942429 PMID: 31545748
[106] Yang HY, Staib LH. Dual Adversarial Autoencoder for Dermoscopy image Generative Modeling. Int Sympos Biomed Imag 2019; 2019: 1247-50.
[107] Abdelhalim ISA, Mohamed MF, Mahdy YB. Data augmentation for skin lesion using self-attention based progressive generative adversarial network. Expert Syst Appl 2021; 165: 113922. http://dx.doi.org/10.1016/j.eswa.2020.113922
[108] Afza F, Khan MA, Sharif M, *et al.* Skin lesion classification: An optimized framework of optimal color features selection. Int Conf Comput Inform Sci (ICCIS) 2020; 2020: 1-6.
[109] Mporas I, Perikos I, Paraskevas M. Color models for skin lesion classification from dermoscopy images//Advances in Integrations of Intelligent Methods. Singapore: Springer 2020; pp. 85-98.
[110] Monisha M, Suresh A, Bapu BRT, Rashmi MR. Classification of malignant melanoma and benign skin lesion by using back propagation neural network and ABCD rule. Cluster Comput 2019; 22(5): 12897-907. http://dx.doi.org/10.1007/s10586-018-1798-7
[111] Chatterjee S, Dey D, Munshi S, Gorai S. Dermatological expert system implementing the ABCD rule of dermoscopy for skin disease identification. Expert Syst Appl 2021; 167: 114204. http://dx.doi.org/10.1016/j.eswa.2020.114204
[112] Yang J, Sun X, Liang J, *et al.* Clinical skin lesion diagnosis using representations inspired by dermatologist criteria. IEEE/CVF Conf Comput Vision Pattern Recogn (CVPR) 2018; 2018: 18311822.
[113] Dhivyaa CR, Sangeetha K, Balamurugan M, Amaran S, Vetriselvi T, Johnpaul P. Skin lesion classification using decision trees and random forest algorithms. J Ambient Intell Humaniz Comput 2020; 2020: 1-13. http://dx.doi.org/10.1007/s12652-020-02675-8
[114] Milton MAA. Automated skin lesion classification using ensemble of deep neural networks in ISIC 2018: Skin lesion analysis towards melanoma detection challenge. arXiv 2019; 2019: 1901.10802.
[115] Singhal A, Shukla R, Kankar PK, Dubey S, Singh S, Pachori RB. Comparing the capabilities of transfer learning models to detect skin lesion in humans. Proc Inst Mech Eng H 2020; 234(10): 1083-93. http://dx.doi.org/10.1177/0954411920939829 PMID: 32643539
[116] Polevaya T, Ravodin R, Filchenkov A. Skin lesion primary morphology classification with end-to-end deep learning network. Int Conf Artif Intell Inform Commun (ICAIIC) 2019; 2019: 247-50. http://dx.doi.org/10.1109/ICAIIC.2019.8668980
[117] Shin HC, Roth HR, Gao M, *et al.* Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Trans Med Imaging 2016; 35(5): 1285-98. http://dx.doi.org/10.1109/TMI.2016.2528162 PMID: 26886976
[118] Qin Z, Liu Z, Zhu P, Xue Y. A GAN-based image synthesis method for skin lesion classification. Comput Methods Programs Biomed 2020; 195: 105568. http://dx.doi.org/10.1016/j.cmpb.2020.105568 PMID: 32526536
[119] Deng J, Dong W, Socher R, *et al.* Imagenet: A large-scale hierarchical image database. IEEE Conf Comput Vision Pattern Recogn 2009; 2009: 248-55.
[120] Jaworek-Korjakowska J, Kleczek P, Gorgon M. Melanoma thickness prediction based on convolutional neural network with VGG-19 model transfer learning. Proc IEEE/CVF Conf Comput Vision Pattern Recogn Workshops 2019; 2019: 00333. http://dx.doi.org/10.1109/CVPRW.2019.00333
[121] Hekler A, Utikal JS, Enk AH, *et al.* Pathologist-level classification of histopathological melanoma images with deep neural networks. Eur J Cancer 2019; 115: 79-83. http://dx.doi.org/10.1016/j.ejca.2019.04.021 PMID: 31129383
[122] Kwasigroch A, Grochowski M, Mikołajczyk A. Neural architecture search for skin lesion classification. IEEE Access 2020; 8: 9061-71. http://dx.doi.org/10.1109/ACCESS.2020.2964424
[123] Brinker TJ, Hekler A, Enk AH, *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: 148-54. http://dx.doi.org/10.1016/j.ejca.2019.02.005 PMID: 30852421
[124] Muckatira S. Properties of winning tickets on skin lesion classification. arXiv 2020; 2020: 1901.10802.
[125] Ratul M A R, Mozaffari MH, Lee WS, *et al.* Skin lesions classification using deep learning based on dilated convolution BioRxiv 2020; 860700.
[126] Tschandl P, Argenziano G, Razmara M, Yap J. Diagnostic accuracy of content-based dermatoscopic image retrieval with deep classification features. Br J Dermatol 2019; 181(1): 155-65. http://dx.doi.org/10.1111/bjd.17189 PMID: 30207594
[127] Allegretti S, Bolelli F, Pollastri F, *et al.* Supporting skin lesion diagnosis with content-based image retrieval. Int Conf Pattern Recogn (ICPR) 2020; 2020: 20591924.
[128] Barata C, Celebi ME, Marques JS. Explainable skin lesion diagnosis using taxonomies. Pattern Recognit 2021; 110: 107413. http://dx.doi.org/10.1016/j.patcog.2020.107413
[129] Barata C, Marques JS, Emre Celebi M. Deep attention model for the hierarchical diagnosis of skin lesions. Proc IEEE/CVF Conf Comput Vision Pattern Recogn Workshops 2019; 2019: 00334. http://dx.doi.org/10.1109/CVPRW.2019.00334
[130] Aggarwal A, Das N, Sreedevi I. Attention-guided deep convolutional neural networks for skin cancer classification. Int Conf Image Proc Theory Tools Appl (IPTA) 2019; 2019: 1-6.
[131] Zhang J, Xie Y, Xia Y, Shen C. Attention residual learning for skin lesion classification. IEEE Trans Med Imaging 2019; 38(9): 2092-103. http://dx.doi.org/10.1109/TMI.2019.2893944 PMID: 30668469
[132] Zhang H, Wu C, Zhang Z, *et al.* Resnest: Split-attention networks. arXiv 2020; 2020: 2004.08955.
[133] Lee I, Kim D, Kang S, *et al.* Ensemble deep learning for skeleton-based action recognition using temporal sliding lstm networks. Proc IEEE Int Conf Comput Vision 2017; 2017: 1012-20. http://dx.doi.org/10.1109/ICCV.2017.115
[134] Wang W, Sun G. Classification and research of skin lesions based on machine learning computers. Mater Cont 2020; 62(3): 1187-200. http://dx.doi.org/10.32604/cmc.2020.05883
[135] Mahbod A, Schaefer G, Ellinger I, Ecker R, Pitiot A, Wang C. Fusing fine-tuned deep features for skin lesion classification. Comput Med Imaging Graph 2019; 71: 19-29. http://dx.doi.org/10.1016/j.compmedimag.2018.10.007 PMID: 30458354
[136] Perez F, Avila S, Valle E. Solo or ensemble? choosing a cnn architecture for melanoma classification. Proc IEEE/CVF Conf Comput Vision Pattern Recogn Workshops 2019; 2019: 1904.12724.
[137] Harangi B, Baran A, Hajdu A. Assisted deep learning framework for multi-class skin lesion classification considering a binary classification support. Biomed Signal Process Control 2020; 62: 102041. http://dx.doi.org/10.1016/j.bspc.2020.102041
[138] Hameed N, Shabut AM, Ghosh MK, Hossain MA. Multi-class multi-level classification algorithm for skin lesions classification

using machine learning techniques. Expert Syst Appl 2020; 141: 112961.
http://dx.doi.org/10.1016/j.eswa.2019.112961

[139] Mahbod A, Schaefer G, Wang C, Dorffner G, Ecker R, Ellinger I. Transfer learning using a multi-scale and multi-network ensemble for skin lesion classification. Comput Methods Programs Biomed 2020; 193: 105475.
http://dx.doi.org/10.1016/j.cmpb.2020.105475 PMID: 32268255

[140] Tang P, Liang Q, Yan X, Xiang S, Zhang D. GP-CNN-DTEL: Global-part CNN model with data-transformed ensemble learning for skin lesion classification. IEEE J Biomed Health Inform 2020; 24(10): 2870-82.
http://dx.doi.org/10.1109/JBHI.2020.2977013 PMID: 32142460

[141] Ghalejoogh GS, Kordy HM, Ebrahimi F. A hierarchical structure based on Stacking approach for skin lesion classification. Expert Syst Appl 2020; 145: 113127.
http://dx.doi.org/10.1016/j.eswa.2019.113127

[142] Walker BN, Rehg JM, Kalra A, *et al.* Dermoscopy diagnosis of cancerous lesions utilizing dual deep learning algorithms *via* visual and audio (sonification) outputs: Laboratory and prospective observational studies. EBioMedicine 2019; 40: 176-83.
http://dx.doi.org/10.1016/j.ebiom.2019.01.028 PMID: 30674442

[143] Sabbaghi S, Aldeen M, Garnavi R. A deep bag-of-features model for the classification of melanomas in dermoscopy images. Ann Int Conf IEEE Eng Med Biol Soc (EMBC) 2016; 2016: 1369-72.

[144] Ahmad B, Usama M, Huang C M, *et al.* Discriminative feature learning for skin disease classification using deep convolutional neural network. IEEE Access 2020; PP(99): 1-1.

[145] Lin T Y, Goyal P, Girshick R, *et al.* Focal loss for dense object detection. IEEE Trans Pattern Anal Machine Intell 2017; PP(99): 2999-3007.

[146] Goceri E. Analysis of deep networks with residual blocks and different activation functions: Classification of skin diseases. Int Conf Image Proc Theory Tools Appl (IPTA) 2019; 2019: 1-6.

[147] Shi X, Dou Q, Xue C, *et al.* An active learning approach for reducing annotation cost in skin lesion analysis. Int Workshop Machine Learn Medical Imag 2019; 2019: 628-36.

[148] Bdair T, Navab N, Albarqouni S. Peer learning for skin lesion classification. arXiv 2021; 2021: 2103.03703.

[149] Bagchi S, Banerjee A, Bathula DR. Learning a meta-ensemble technique for skin lesion classification and novel class detection. Proc IEEE/CVF Conf Comput Vision Pattern Recogn Workshops 2020; 2020: 746-7.
http://dx.doi.org/10.1109/CVPRW50498.2020.00381

[150] Combalia M, Hueto F, Puig S, *et al.* Uncertainty estimation in deep neural networks for dermoscopy image classification. Proc IEEE/CVF Conf Comput Vision Pattern Recogn Workshops 2020; 2020: 744-5.

[151] Jinnai S, Yamazaki N, Hirano Y, Sugawara Y, Ohe Y, Hamamoto R. The development of a skin cancer classification system for pigmented skin lesions using deep learning. Biomolecules 2020; 10(8): 1123.
http://dx.doi.org/10.3390/biom10081123 PMID: 32751349

[152] Khamparia A, Singh PK, Rani P, *et al.* An internet of health things-driven deep learning framework for detection and classification of skin cancer using transfer learning. Trans Emerg Telecommun Technol 2020; 2020: e3963.

[153] Hameed N, Shabut A, Hameed F, *et al.* An intelligent inflammatory skin lesions classification scheme for mobile devices. Int Conf Comput Electron Commun Eng (iCCECE) 2019; 2019: 83-8.

[154] Weingast J, Scheibböck C, Wurm EMT, *et al.* A prospective study of mobile phones for dermatology in a clinical setting. J Telemed Telecare 2013; 19(4): 213-8.
http://dx.doi.org/10.1177/1357633x13490890 PMID: 24163062

[155] Hogan K, Cullan J, Patel V, Rajpara A, Aires D. Overcalling a teledermatology selfie: A new twist in a growing field. Dermatol Online J 2015; 21(6): 13030/qt84x5d2gg.
http://dx.doi.org/10.5070/D3216027826 PMID: 26158371

[156] Ge Z, Demyanov S, Chakravorty R, *et al.* Skin disease recognition using deep saliency features and multimodal learning of dermoscopy and clinical images. Int Conf Med Image Comput Comput-Assist Interven 2017; 2017: 250-8.

[157] 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 2018; 23(2): 538-46.
http://dx.doi.org/10.1109/JBHI.2018.2824327 PMID: 29994053

[158] Nunnari F, Bhuvaneshwara C, Ezema AO, *et al.* A study on the fusion of pixels and patient metadata in CNN-based classification of skin lesion images. Int Cross-Domain Conf Machine Learn Knowledge Extract 2020; 2020: 1-17.

[159] Yap J, Yolland W, Tschandl P. Multimodal skin lesion classification using deep learning. Exp Dermatol 2018; 27(11): 1261-7.
http://dx.doi.org/10.1111/exd.13777 PMID: 30187575

[160] Pacheco AGC, Krohling RA. The impact of patient clinical information on automated skin cancer detection. Comput Biol Med 2020; 116: 103545.
http://dx.doi.org/10.1016/j.compbiomed.2019.103545 PMID: 31760271

[161] Bi L, Feng DD, Fulham M, Kim J. Multi-Label classification of multi-modality skin lesion *via* hyper-connected convolutional neural network. Pattern Recognit 2020; 107: 107502.
http://dx.doi.org/10.1016/j.patcog.2020.107502

[162] Razmjooy N, Ashourian M, Karimifard M, *et al.* Computer-aided diagnosis of skin cancer: A review. Curr Med Imaging 2020; 16(7): 781-93.
http://dx.doi.org/10.2174/1573405616666200129095242

[163] Al Mamun M, Uddin MS. A comparative study among segmentation techniques for skin disease detection systems. Proc Int Conf Trends Comput Cogn Eng 2021; 2021: 155-67.
http://dx.doi.org/10.1007/978-981-33-4673-4_14

[164] Celebi M E, Wen Q, Iyatomi H, *et al.* A state-of-the-art survey on lesion border detection in dermoscopy images. Dermoscopy Image Anal 2015; 10: 97-129.
http://dx.doi.org/10.1201/b19107-8

[165] Pathan S, Prabhu KG, Siddalingaswamy PC. Techniques and algorithms for computer aided diagnosis of pigmented skin lesions—A review. Biomed Signal Process Control 2018; 39: 237-62.
http://dx.doi.org/10.1016/j.bspc.2017.07.010

[166] Chang H. Skin cancer reorganization and classification with deep neural network. arXiv 2017; 2017: 1703.00534.

[167] Rashid Sheykhahmad F, Razmjooy N, Ramezani M. A novel method for skin lesion segmentation. Int J Inform Secur Sys Manage 2015; 4(2): 458-66.

[168] Ali AR, Li J, O'Shea SJ, *et al.* A deep learning based approach to skin lesion border extraction with a novel edge detector in dermoscopy images. Int Joint Conf Neural Networks (IJCNN) 2019; 2019: 1-7.

[169] Jayalakshmi D, Dheeba J. Border detection in skin lesion images using an improved clustering algorithm. Int J e-Collaborat (IJeC) 2020; 16(4): 15-29.

[170] Sengupta S, Mittal N, Modi M. Improved skin lesion edge detection method using Ant Colony Optimization. Skin Res Technol 2019; 25(6): 846-56.
http://dx.doi.org/10.1111/srt.12744 PMID: 31228313

[171] Abbas A A, Abu-Almash F S. Skin lesion border detection based on optimal statistical model using optimized colour channel. J Autonom Intell 2020; 3(1): 18-26.

[172] Bayraktar M, Kockara S, Halic T, Mete M, Wong HK, Iqbal K. Local edge-enhanced active contour for accurate skin lesion border detection. BMC Bioinform 2019; 20(Suppl. 2): 91.
http://dx.doi.org/10.1186/s12859-019-2625-8 PMID: 30871471

[173] Abeysinghe D, Sotheeswaran S. Novel computational approaches for border irregularity prediction to detect melanoma in skin lesions. Int Res Conf Smart Comput Sys Eng (SCSE) 2020; 2020: 216-22.

[174] Han SS, Park GH, Lim W, *et al.* Deep neural networks show an equivalent and often superior performance to dermatologists in onychomycosis diagnosis: Automatic construction of onychomycosis datasets by region-based convolutional deep neural network. PLoS One 2018; 13(1): e0191493.
http://dx.doi.org/10.1371/journal.pone.0191493 PMID: 29352285

[175] Ali AR, Li J, Yang G, O'Shea SJ. A machine learning approach to automatic detection of irregularity in skin lesion border using dermoscopic images. PeerJ Comput Sci 2020; 6: e268.
http://dx.doi.org/10.7717/peerj-cs.268 PMID: 33816919

[176] Ali AR, Li J, Kanwal S, *et al.* A novel fuzzy multilayer perceptron (F-MLP) for the detection of irregularity in skin lesion border using dermoscopy images. Front Med 2020; 2020: 7.

[177] Zhang G, Hsu CHR, Lai H, Zheng X. Deep learning based feature representation for automated skin histopathological image annotation. Multimedia Tools Appl 2018; 77(8): 9849-69. http://dx.doi.org/10.1007/s11042-017-4788-5

[178] Bozkurt A, Kose K, Alessi-Fox C, *et al.* A multiresolution convolutional neural network with partial label training for annotating reflectance confocal microscopy images of skin. Int Conf Med Image Comput Comput-Assist Int 2018; 2018: 1802.02213.

[179] Goyal M, Yap MH, Hassanpour S. Multi-class semantic segmentation of skin lesions *via* fully convolutional networks. arXiv 2017; 2017: 1711.10449.

[180] Liu Z, Pan H, Gong C, *et al.* Multi-class skin lesion segmentation for cutaneous T-cell lymphomas on high-resolution clinical images. Int Conf Med Image Comput Comput-Assist Interven 2020; 2020: 351-61.

[181] Moradi N, Mahdavi-Amiri N. Multi-class segmentation of skin lesions *via* joint dictionary learning. Biomed Signal Process Control 2021; 68: 102787. http://dx.doi.org/10.1016/j.bspc.2021.102787

[182] Moradi N, Mahdavi-Amiri N. Kernel sparse representation based model for skin lesions segmentation and classification. Comput Methods Programs Biomed 2019; 182: 105038. http://dx.doi.org/10.1016/j.cmpb.2019.105038 PMID: 31437709

[183] Delong A, Osokin A, Isack HN, Boykov Y. Fast approximate energy minimization with label costs. Int J Comput Vis 2012; 96(1): 1-27. http://dx.doi.org/10.1007/s11263-011-0437-z

[184] Thomas SM, Lefevre JG, Baxter G, Hamilton NA. Interpretable deep learning systems for multi-class segmentation and classification of non-melanoma skin cancer. Med Image Anal 2021; 68: 101915. http://dx.doi.org/10.1016/j.media.2020.101915 PMID: 33260112

[185] Garnavi R, Aldeen M, Celebi ME, *et al.* Automatic segmentation of dermoscopy images using histogram thresholding on optimal color channels. Int J Med Med Sci 2010; 1(2): 126-34.

[186] Salih O, Viriri S. Skin lesion segmentation using stochastic region-merging and pixel-based Markov random field. Symmetry (Basel) 2020; 12(8): 1224. http://dx.doi.org/10.3390/sym12081224

[187] Rizzi M, Guaragnella C. Skin lesion segmentation using image bit-plane multilayer approach. Appl Sci (Basel) 2020; 10(9): 3045. http://dx.doi.org/10.3390/app10093045

[188] Razmjooy N, Mousavi BS, Soleymani F, *et al.* A computer-aided diagnosis system for malignant melanomas. Neural Comput Appl 2013; 23(7): 2059-71. http://dx.doi.org/10.1007/s00521-012-1149-1

[189] Patiño D, Avendaño J, Branch JW. Automatic skin lesion segmentation on dermoscopy images by the means of superpixel merging. Int Conf Med Image Comput Comput-Assist Interven 2018; 2018: 728-36.

[190] Filali I, Belkadi M. Multi-scale contrast based skin lesion segmentation in digital images. Optik (Stuttg) 2019; 185: 794-811. http://dx.doi.org/10.1016/j.ijleo.2019.04.022

[191] Devi SS, Singh NH, Laskar RH. Fuzzy C-means clustering with histogram based cluster selection for skin lesion segmentation using non-dermoscopy images. Int J Interact Multimedia Artif Intell 2020; 6(1): 26-31.

[192] Peruch F, Bogo F, Bonazza M, Cappelleri VM, Peserico E. Simpler, faster, more accurate melanocytic lesion segmentation through MEDS. IEEE Trans Biomed Eng 2014; 61(2): 557-65. http://dx.doi.org/10.1109/TBME.2013.2283803 PMID: 24081839

[193] Ma Z, Tavares JMRS. A novel approach to segment skin lesions in dermoscopy images based on a deformable model. IEEE J Biomed Health Inform 2016; 20(2): 615-23. http://dx.doi.org/10.1109/JBHI.2015.2390032 PMID: 25585429

[194] Pereira PMM, Fonseca-Pinto R, Paiva RP, *et al.* Skin lesion classification enhancement using border-line features–The melanoma *vs.* nevus problem. Biomed Signal Process Control 2020; 57: 101765. http://dx.doi.org/10.1016/j.bspc.2019.101765

[195] Hasan MJ, Uddin J, Pinku SN. A novel modified SFTA approach for feature extraction. Int Conf Electrical Eng Inf Commun Technol (ICEEICT) 2016; 2016: 1-5.

[196] Parida P, Rout R. Transition region based approach for skin lesion segmentation. ELCVIA 2020; 19(1): 28-37. http://dx.doi.org/10.5565/rev/elcvia.1177

[197] Ruela M, Barata C, Marques JS, Rozeira J. A system for the detection of melanomas in dermoscopy images using shape and symmetry features. Comput Methods Biomech Biomed Eng Imaging Vis 2017; 5(2): 127-37. http://dx.doi.org/10.1080/21681163.2015.1029080

[198] Nasir M, Attique Khan M, Sharif M, Lali IU, Saba T, Iqbal T. An improved strategy for skin lesion detection and classification using uniform segmentation and feature selection based approach. Microsc Res Tech 2018; 81(6): 528-43. http://dx.doi.org/10.1002/jemt.23009 PMID: 29464868

[199] Asaeikheybari G, Green J, Qian X, Jiang H, Huang M-C. Medical image learning from a few/few training samples: Melanoma segmentation study. Smart Health (Amst) 2019; 14: 100088. http://dx.doi.org/10.1016/j.smhl.2019.100088

[200] McIntosh LM, Mansfield JR, Crowson AN, Mantsch HH, Jackson M. Analysis and interpretation of infrared microscopic maps: Visualization and classification of skin components by digital staining and multivariate analysis. Biospectroscopy 1999; 5(5): 265-75. http://dx.doi.org/10.1002/(SICI)1520-6343(1999)5:5<265::AID-BSPY1>3.0.CO;2-F

[201] McIntosh LM, Summers R, Jackson M, *et al.* Towards non-invasive screening of skin lesions by near-infrared spectroscopy. J Invest Dermatol 2001; 116(1): 175-81. http://dx.doi.org/10.1046/j.1523-1747.2001.00212.x PMID: 11168814

[202] Mishra R, Daescu O. Deep learning for skin lesion segmentation. IEEE Int Conf Bioinform Biomed (BIBM) 2017; 2017: 1189-94.

[203] Zhang X. Melanoma segmentation based on deep learning. CAS 2017; 22(sup1): 267-77. http://dx.doi.org/10.1080/24699322.2017.1389405

[204] Peng Y, Wang N, Wang Y, Wang M. Segmentation of dermoscopy image using adversarial networks. Multimedia Tools Appl 2019; 78(8): 10965-81. http://dx.doi.org/10.1007/s11042-018-6523-2

[205] Kaymak R, Kaymak C, Ucar A. Skin lesion segmentation using fully convolutional networks: A comparative experimental study. Expert Syst Appl 2020; 161: 113742. http://dx.doi.org/10.1016/j.eswa.2020.113742

[206] Öztürk Ş, Özkaya U. Skin lesion segmentation with improved convolutional neural network. J Digit Imaging 2020; 33(4): 958-70. http://dx.doi.org/10.1007/s10278-020-00343-z PMID: 32378058

[207] He K, Zhang X, Ren S, *et al.* Deep residual learning for image recognition. Proc IEEE Conf Comput Vision Pattern Recogn 2016; 2016: 770-8.

[208] Huang G, Liu Z, Van Der Maaten L, *et al.* Densely connected convolutional networks. Proc IEEE Conf Comput Vision Pattern Recogn 2017; 2017: 4700-8.

[209] Nasr-Esfahani E, Rafiei S, Jafari MH, *et al.* Dense pooling layers in fully convolutional network for skin lesion segmentation. Comput Med Imaging Graph 2019; 78: 101658. http://dx.doi.org/10.1016/j.compmedimag.2019.101658 PMID: 31634739

[210] Wei Z, Song H, Chen L, Li Q, Han G. Attention-based DenseUnet network with adversarial training for skin lesion segmentation. IEEE Access 2019; 7: 136616-29. http://dx.doi.org/10.1109/ACCESS.2019.2940794

[211] Jiang F, Zhou F, Qin J, *et al.* Decision-augmented generative adversarial network for skin lesion segmentation. Int Sympos Biomed Imag 2019; 2019: 447-50.

[212] Bi L, Feng D, Fulham M, *et al.* Improving skin lesion segmentation *via* stacked adversarial learning. Int Sympos Biomed Imag 2019; 2019: 1100-3.

[213] Tu W, Liu X, Hu W, *et al.* Segmentation of lesion in dermoscopy images using dense-residual network with adversarial learning. IEEE Int Conf Image Proc (ICIP) 2019; 2019: 1430-4.

[214] Lei B, Xia Z, Jiang F, *et al.* Skin lesion segmentation *via* generative adversarial networks with dual discriminators. Med Image Anal 2020; 64: 101716.
http://dx.doi.org/10.1016/j.media.2020.101716 PMID: 32492581

[215] Tschandl P, Sinz C, Kittler H. Domain-specific classification-pretrained fully convolutional network encoders for skin lesion segmentation. Comput Biol Med 2019; 104: 111-6.
http://dx.doi.org/10.1016/j.compbiomed.2018.11.010 PMID: 30471461

[216] Chaurasia A, Culurciello E. Linknet: Exploiting encoder representations for efficient semantic segmentation. IEEE Visual Commun Image Proc (VCIP) 2017; 2017: 1-4.

[217] Soudani A, Barhoumi W. An image-based segmentation recommender using crowdsourcing and transfer learning for skin lesion extraction. Expert Syst Appl 2019; 118: 400-10.
http://dx.doi.org/10.1016/j.eswa.2018.10.029

[218] Phillips A, Teo I, Lang J. Segmentation of prognostic tissue structures in cutaneous melanoma using whole slide images. Proc IEEE/CVF Conf Comput Vision Pattern Recogn Workshops 2019; 2019: 00332.
http://dx.doi.org/10.1109/CVPRW.2019.00332

[219] PascalVOC [EB/OL]. Available from: http://hostrobotsoxacuk/pascal/VOC/

[220] Canalini L, Pollastri F, Bolelli F, *et al.* Skin lesion segmentation ensemble with diverse training strategies. Int Conf Comput Anal Images Patterns 2019; 2019: 89-101.

[221] Bagheri F, Tarokh MJ, Ziaratban M. Skin lesion segmentation from dermoscopy images by using Mask R-CNN, Retina-Deeplab, and graph-based methods. Biomed Signal Process Control 2021; 67: 102533.
http://dx.doi.org/10.1016/j.bspc.2021.102533

[222] Hasan MK, Elahi MTE, Alam MA, *et al.* DermoExpert: Skin lesion classification using a hybrid convolutional neural network through segmentation, transfer learning, and augmentation. medRxiv 2021; 2021: 21251038.
http://dx.doi.org/10.1101/2021.02.02.21251038

[223] Xiao J, Xu H, Zhao W, Cheng C, Gao HH. A Prior-mask-guided Few-shot Learning for Skin Lesion Segmentation. Computing 2021; 2021: 1-23.
http://dx.doi.org/10.1007/s00607-021-00907-z

[224] Jin FQ, Knight AE, Cardones AR, Nightingale KR, Palmeri ML. Semi-automated weak annotation for deep neural network skin thickness measurement. Ultrason Imaging 2021; 43(4): 167-74.
http://dx.doi.org/10.1177/01617346211014138 PMID: 33971769

[225] Messadi M, Cherifi H, Bessaid A. Segmentation and ABCD rule extraction for skin tumors classification. arXiv 2021; 2021: 2106.04372.

[226] Lin BS, Michael K, Kalra S, *et al.* Skin lesion segmentation: U-Nets *versus* clustering. IEEE Sympos Series Comput Intell (SSCI) 2017; 2017: 1-7.

[227] Huang C, Yu Y. Skin lesion segmentation based on deep learning. Int Conf Commun Technol (ICCT) 2020; 2020: 1360-4.

[228] Justin S, Pattnaik M. Skin lesion segmentation by pixel by pixel approach using deep learning. IJASIS 2020; 6(1): 12-20.

[229] Zafar K, Gilani SO, Waris A, *et al.* Skin lesion segmentation from dermoscopy images using convolutional neural network. Sensors (Basel) 2020; 20(6): 1601.
http://dx.doi.org/10.3390/s20061601 PMID: 32183041

[230] Li W, Raj ANJ, Tjahjadi T, *et al.* Digital hair removal by deep learning for skin lesion segmentation. Pattern Recognit 2021; 117: 107994.
http://dx.doi.org/10.1016/j.patcog.2021.107994

[231] Ramya J, Vijaylakshmi HC, Saifuddin HM. Segmentation of skin lesion images using discrete wavelet transform. Biomed Signal Process Control 2021; 69: 102839.
http://dx.doi.org/10.1016/j.bspc.2021.102839

[232] Dastane T, Rao V, Shenoy K, *et al.* An effective pixel-wise approach for skin colour segmentation using pixel neighbourhood technique. arXiv 2021; 2021: 2108.10971.

[233] Filali I, Belkadi M, Aoudjit R, Lalam M. Graph weighting scheme for skin lesion segmentation in macroscopic images. Biomed Signal Process Control 2021; 68: 102710.
http://dx.doi.org/10.1016/j.bspc.2021.102710

[234] Adegun A, Viriri S. Deep convolutional network-based framework for melanoma lesion detection and segmentation. Int Conf Adv Concepts Intell Vision Sys 2020; 2020: 51-62.

[235] Xu Z, Sheykhahmad FR, Ghadimi N, Razmjooy N. Computer-aided diagnosis of skin cancer based on soft computing techniques. Open Med (Wars) 2020; 15(1): 860-71.
http://dx.doi.org/10.1515/med-2020-0131 PMID: 33336044

[236] Razmjooy N, Sheykhahmad FR, Ghadimi N. A hybrid neural network–world cup optimization algorithm for melanoma detection. Open Med (Wars) 2018; 13(1): 9-16.
http://dx.doi.org/10.1515/med-2018-0002 PMID: 29577090

[237] Razmjooy N, Mousavi BS, Soleymani F. A hybrid neural network Imperialist Competitive Algorithm for skin color segmentation. Math Comput Model 2013; 57(3-4): 848-56.
http://dx.doi.org/10.1016/j.mcm.2012.09.013

[238] Adegun AA, Viriri S, Yousaf MH. A Probabilistic-based deep learning model for skin lesion segmentation. Appl Sci (Basel) 2021; 11(7): 3025.
http://dx.doi.org/10.3390/app11073025

[239] Qiu Y, Cai J, Qin X, *et al.* Inferring skin lesion segmentation with fully connected CRFS based on multiple deep convolutional neural networks. IEEE Access 2020; 8: 144246-58.

[240] Khan MA, Sharif M, Akram T, Damaševičius R, Maskeliūnas R. Skin lesion segmentation and multiclass classification using deep learning features and improved moth flame optimization. Diagnostics (Basel) 2021; 11(5): 811.
http://dx.doi.org/10.3390/diagnostics11050811 PMID: 33947117

[241] Shan P, Wang Y, Fu C, Song W, Chen J. Automatic skin lesion segmentation based on FC-DPN. Comput Biol Med 2020; 123: 103762.
http://dx.doi.org/10.1016/j.compbiomed.2020.103762 PMID: 32768035

[242] Jiang C, Zhang Y, Wang J, *et al.* Approximated masked global context network for skin lesion segmentation. Int Conf Artif Neural Networks 2021; 2021: 610-22.

[243] Qamar S, Ahmad P, Shen L. Dense Encoder-Decoder–Based Architecture for Skin Lesion Segmentation. Cognit Comput 2021; 13(2): 583-94.
http://dx.doi.org/10.1007/s12559-020-09805-6

[244] Chen LC, Papandreou G, Schroff F, Adam H. Rethinking atrous convolution for semantic image segmentation. arXiv 2017; 2017: 1706.05587.

[245] Xie F, Yang J, Liu J, Jiang Z, Zheng Y, Wang Y. Skin lesion segmentation using high-resolution convolutional neural network. Comput Methods Programs Biomed 2020; 186: 105241.
http://dx.doi.org/10.1016/j.cmpb.2019.105241 PMID: 31837637

[246] Sarker M, Kamal M, Rashwan HA, *et al.* MobileGAN: Skin lesion segmentation using a lightweight generative adversarial network. arXiv 2019; 2019: 1907.00856.

[247] Jiang Y, Cao S, Tao S, Zhang H. Skin lesion segmentation based on multi-scale attention convolutional neural network. IEEE Access 2020; 8: 122811-25.
http://dx.doi.org/10.1109/ACCESS.2020.3007512

[248] Oliveira RB, Pereira AS, Tavares JMR. S Computational diagnosis of skin lesions from dermoscopy images using combined features. Neural Comput Appl 2019; 31(10): 6091-111.
http://dx.doi.org/10.1007/s00521-018-3439-8

[249] Tong X, Wei J, Sun B, Su S, Zuo Z, Wu P. ASCU-Net: Attention gate, spatial and channel attention u-net for skin lesion segmentation. Diagnostics (Basel) 2021; 11(3): 501.
http://dx.doi.org/10.3390/diagnostics11030501 PMID: 33809048

[250] Arora R, Raman B, Nayyar K, Awasthi R. Automated skin lesion segmentation using attention-based deep convolutional neural network. Biomed Signal Process Control 2021; 65: 102358.
http://dx.doi.org/10.1016/j.bspc.2020.102358

[251] Ren Y, Yu L, Tian S, Cheng J, Guo Z, Zhang Y. Serial attention network for skin lesion segmentation. J Ambient Intell Humaniz Comput 2021; 2021: 1-12.
http://dx.doi.org/10.1007/s12652-021-02933-3

[252] Codella N C F, Nguyen Q B, Pankanti S, *et al.* Deep learning ensembles for melanoma recognition in dermoscopy images. IBM J Res Dev 2017; 61(4/5): 15.
http://dx.doi.org/10.1147/JRD.2017.2708299

[253] Kaya U, Fidan M. Parametric and nonparametric correlation ranking based supervised feature selection methods for skin segmentation. J Ambient Intell Humaniz Comput 2021; 2021: 1-13. http://dx.doi.org/10.1007/s12652-021-02936-0

[254] Yuan Y, Lo YC. Improving dermoscopy image segmentation with enhanced convolutional-deconvolutional networks. IEEE J Biomed Health Inform 2019; 23(2): 519-26. http://dx.doi.org/10.1109/JBHI.2017.2787487 PMID: 29990146

[255] Kaur P, Dana KJ, Cula GO, *et al.* Hybrid deep learning for reflectance confocal microscopy skin disease images. Int Conf Pattern Recogn (ICPR) 2016; 2016: 1466-71.

[256] Pour MP, Seker H. Transform domain representation-driven convolutional neural networks for skin lesion segmentation. Expert Syst Appl 2020; 144: 113129. http://dx.doi.org/10.1016/j.eswa.2019.113129

[257] Abhishek K, Hamarneh G, Drew MS. Illumination-based transformations improve skin lesion segmentation in dermoscopy images. Proc IEEE/CVF Conf Comput Vision Pattern Recogn Workshops 2020; 2020: 728-9.

[258] Attia M, Hossny M, Nahavandi S, *et al.* Skin melanoma segmentation using recurrent and convolutional neural networks. Int Sympos Biomed Imag 2017; 2017: 292-6.

[259] Khatibi T, Rezaei N, Ataei Fashtami L, Totonchi M. Proposing a novel unsupervised stack ensemble of deep and conventional image segmentation (SEDCIS) method for localizing vitiligo lesions in skin images. Skin Res Technol 2021; 27(2): 126-37. http://dx.doi.org/10.1111/srt.12920 PMID: 32662570

[260] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997; 9(8): 1735-80. http://dx.doi.org/10.1162/neco.1997.9.8.1735 PMID: 9377276

[261] Bi L, Kim J, Ahn E, Kumar A, Fulham M, Feng D. Dermoscopy image segmentation *via* multistage fully convolutional networks. IEEE Trans Biomed Eng 2017; 64(9): 2065-74. http://dx.doi.org/10.1109/TBME.2017.2712771 PMID: 28600236

[262] Li H, He X, Zhou F, *et al.* Dense deconvolutional network for skin lesion segmentation. IEEE J Biomed Health Inform 2019; 23(2): 527-37. http://dx.doi.org/10.1109/JBHI.2018.2859898 PMID: 30047917

[263] Li H, He X, Yu Z, *et al.* Skin lesion segmentation *via* dense connected deconvolutional network. Int Conf Pattern Recogn (ICPR) 2018; 2018: 671-5.

[264] Ji W, Cai L, Chen W, *et al.* Segmentation of lesions in skin image based on salient object detection with deeply supervised learning. Int Conf Comput Commun (ICCC) 2018; 2018: 1567-73.

[265] Liu L, Mou L, Zhu XX, Mandal M. Automatic skin lesion classification based on mid-level feature learning. Comput Med Imaging Graph 2020; 84: 101765. http://dx.doi.org/10.1016/j.compmedimag.2020.101765 PMID: 32810817

[266] Bozorgtabar B, Ge Z, Chakravorty R, *et al.* Investigating deep side layers for skin lesion segmentation. Int Sympos Biomed Imag 2017; 2017: 256-60.

[267] Nathan S, Kansal P. Lesion net--skin lesion segmentation using coordinate convolution and deep residual units. arXiv 2020; 2020: 2012.14249.

[268] Huang L, Zhao Y, Yang T. Skin lesion segmentation using object scale-oriented fully convolutional neural networks Signal. Signal Image Video Process 2019; 13(3): 431-8. http://dx.doi.org/10.1007/s11760-018-01410-3

[269] Singh VK, Abdel-Nasser M, Rashwan HA, *et al.* FCA-net: Adversarial learning for skin lesion segmentation based on multi-scale features and factorized channel attention. IEEE Access 2019; 7: 130552-65. http://dx.doi.org/10.1109/ACCESS.2019.2940418

[270] Zhu L, Feng S, Zhu W, *et al.* ASNet: An adaptive scale network for skin lesion segmentation in dermoscopy images//Medical Imaging 2020. Biomedical Applications in Molecular, Structural, and Functional Imaging International Society for Optics and Photonics 2020; 11317: 113170W.

[271] Bi L, Kim J, Ahn E, *et al.* Semi-automatic skin lesion segmentation *via* fully convolutional networks. Int Sympos Biomed Imag 2017; 2017: 561-4.

[272] Mirikharaji Z, Hamarneh G. Star shape prior in fully convolutional networks for skin lesion segmentation. Int Conf Med Image Comput Comput-Assist Interven 2018; 2018: 737-45.

[273] Goceri E. Deep learning based classification of facial dermatological disorders. Comput Biol Med 2021; 128: 104118. http://dx.doi.org/10.1016/j.compbiomed.2020.104118 PMID: 33221639

[274] Zhang J, Petitjean C, Ainouz S. Kappa loss for skin lesion segmentation in fully convolutional network. Int Sympos Biomed Imag 2020; 2020: 2001-4.

[275] Abhishek K, Hamarneh G. Matthews correlation coefficient loss for deep convolutional networks: Application to skin lesion segmentation. Int Sympos Biomed Imag 2021; 2021: 225-9.

[276] Hasan MK, Dahal L, Samarakoon PN, Tushar FI, Martí R. DSNet: Automatic dermoscopic skin lesion segmentation. Comput Biol Med 2020; 120: 103738. http://dx.doi.org/10.1016/j.compbiomed.2020.103738 PMID: 32421644

[277] Zhang N, Cai YX, Wang YY, Tian YT, Wang XL, Badami B. Skin cancer diagnosis based on optimized convolutional neural network. Artif Intell Med 2020; 102: 101756. http://dx.doi.org/10.1016/j.artmed.2019.101756 PMID: 31980095

[278] Ribeiro V, Avila S, Valle E. Less is more: Sample selection and label conditioning improve skin lesion segmentation. Proc IEEE/CVF Conf Comput Vision Pattern Recogn Workshops 2020; 2020: 738-9.

[279] Mirikharaji Z, Abhishek K, Izadi S, *et al.* D-LEMA: Deep learning ensembles from multiple annotations--application to skin lesion segmentation. arXiv 2020; 2020: 2012.07206.

[280] Raj R, Londhe ND, Sonawane R. Automatic psoriasis lesion segmentation from raw color images using deep learning. Int Conf Bioinform Biomed (BIBM) 2020; 2020: 723-8.

[281] Udrea A, Mitra GD. Generative adversarial neural networks for pigmented and non-pigmented skin lesions detection in clinical images. In: 2017 21st International Conference on Control Systems and Computer Science (CSCS). 2017 May 29-31; Bucharest, Romania. 364-8.

[282] Pal A, Garain U, Chandra A, Chatterjee R, Senapati S. Psoriasis skin biopsy image segmentation using deep convolutional neural network. Comput Methods Programs Biomed 2018; 159: 59-69. http://dx.doi.org/10.1016/j.cmpb.2018.01.027 PMID: 29650319

[283] Bozkurt A, Gale T, Kose K, *et al.* Delineation of skin strata in reflectance confocal microscopy images with recurrent convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2017 Jul 21-26; Honolulu, HI, USA. 25-33. http://dx.doi.org/10.1109/CVPRW.2017.108

[284] Maglogiannis I, Delibasis KK. Enhancing classification accuracy utilizing globules and dots features in digital dermoscopy. Comput Methods Programs Biomed 2015; 118(2): 124-33. http://dx.doi.org/10.1016/j.cmpb.2014.12.001 PMID: 25540998

[285] Czajkowska J, Badura P, Korzekwa S, Płatkowska-Szczerek A. Deep learning approach to skin layers segmentation in inflammatory dermatoses. Ultrasonics 2021; 114: 106412. http://dx.doi.org/10.1016/j.ultras.2021.106412 PMID: 33784575

[286] Sarker MMK, Rashwan HA, Akram F, *et al.* SLSNet: Skin lesion segmentation using a lightweight generative adversarial network. Expert Syst Appl 2021; 183: 115433. http://dx.doi.org/10.1016/j.eswa.2021.115433

[287] Wibowo A, Purnama SR, Wirawan PW, *et al.* Lightweight encoder-decoder model for automatic skin lesion segmentation. Inform Med Unlocked 2021; 25: 100640.

[288] Premaladha J, Ravichandran KS. Novel approaches for diagnosing melanoma skin lesions through supervised and deep learning algorithms. J Med Syst 2016; 40(4): 96. http://dx.doi.org/10.1007/s10916-016-0460-2 PMID: 26872778

[289] Patiño D, Ceballos-Arroyo AM, Rodriguez-Rodriguez JA, *et al.* Melanoma detection on dermoscopy images using superpixels segmentation and shape-based features. In: Proc SPIE 11330, 15th International Symposium on Medical Information Processing and Analysis. 2019 Nov 6-8; Medelin, Colombia. 1133018.

[290] Aishwarya U, Daniel IJ, Raghul R. Convolutional neural network based skin lesion classification and identification. In: 2020 Inter-

national Conference on Inventive Computation Technologies (ICICT). 2020 Feb 26-28; Coimbatore, India. 264-70.

[291] Sikkandar MY, Alrasheadi BA, Prakash NB, *et al.* Deep learning based an automated skin lesion segmentation and intelligent classification model. J Ambient Intell Humaniz Comput 2021; 12: 3245-55.

[292] Amin J, Sharif A, Gul N, *et al.* Integrated design of deep features fusion for localization and classification of skin cancer. Pattern Recognit Lett 2020; 131: 63-70. http://dx.doi.org/10.1016/j.patrec.2019.11.042

[293] Al Nazi Z, Abir TA. Automatic skin lesion segmentation and melanoma detection: Transfer learning approach with U-Net and DCNN-SVM. In: Proceedings of International Joint Conference on Computational Intelligence. 371-81.

[294] Almaraz-Damian J-A, Ponomaryov V, Sadovnychiy S. Melanoma and nevus skin lesion classification using handcraft and deep learning feature fusion *via* mutual information measures. Entropy 2020; 22: 484.

[295] Prathiba M, Jose D, Saranya R. Automated melanoma recognition in dermoscopy images *via* very deep residual networks. IOP Conf Ser: Mater Sci Eng 2019; 561(1): 12107.

[296] Khan MA, Akram T, Zhang YD, Sharif M. Attributes based skin lesion detection and recognition: A mask RCNN and transfer learning-based deep learning framework. Pattern Recognit Lett 2021; 143: 58-66266. http://dx.doi.org/10.1016/j.patrec.2020.12.015

[297] Jayapriya K, Jacob IJ. Hybrid fully convolutional networks-based skin lesion segmentation and melanoma detection using deep feature. Int J Imaging Syst Technol 2020; 30(2): 348-57. http://dx.doi.org/10.1002/ima.22377

[298] Han SS, Moon IJ, Lim W, *et al.* Keratinocytic skin cancer detection on the face using region-based convolutional neural network. JAMA Dermatol 2020; 156(1): 29-37. http://dx.doi.org/10.1001/jamadermatol.2019.3807 PMID: 31799995

[299] Mahbod A, Tschandl P, Langs G, Ecker R, Ellinger I. The effects of skin lesion segmentation on the performance of dermatoscopic image classification. Comput Methods Programs Biomed 2020; 197: 105725. http://dx.doi.org/10.1016/j.cmpb.2020.105725 PMID: 32882594

[300] Maron RC, Hekler A, Krieghoff-Henning E, *et al.* Reducing the Impact of Confounding Factors on Skin Cancer Classification *via* Image Segmentation: Technical Model Study. J Med Internet Res 2021; 23(3): e21695. http://dx.doi.org/10.2196/21695 PMID: 33764307

[301] Al-Masni MA, Kim DH, Kim TS. Multiple skin lesions diagnostics *via* integrated deep convolutional networks for segmentation and classification. Comput Methods Programs Biomed 2020; 190: 105351. http://dx.doi.org/10.1016/j.cmpb.2020.105351 PMID: 32028084

[302] Xie Y, Zhang J, Xia Y, Shen C. A mutual bootstrapping model for automated skin lesion segmentation and classification. IEEE Trans Med Imaging 2020; 39(7): 2482-93. http://dx.doi.org/10.1109/TMI.2020.2972964 PMID: 32070946

[303] Pal A, Chaturvedi A, Garain U, *et al.* Severity grading of psoriatic plaques using deep CNN based multi-task learning. 2016 23rd International Conference on Pattern Recognition (ICPR). 2016 Dec 4-8; Cancan, Mexico. 1478-83.

[304] Vesal S, Patil SM, Ravikumar N, *et al.* A multi-task framework for skin lesion detection and segmentation. In: Stoyanov D, Taylor Z, Sarikaya D, Eds. Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin disease image Analysis. Cham: Springer 2018; pp. 285-93.

[305] Yang X, Zeng Z, Yeo SY, *et al.* A novel multi-task deep learning model for skin lesion segmentation and classification. arXiv 2017.

[306] Li Y, Shen L. Skin lesion analysis towards melanoma detection using deep learning network. Sensors (Basel) 2018; 18(2): 556. http://dx.doi.org/10.3390/s18020556 PMID: 29439500

[307] Song L, Lin J, Wang ZJ, Wang H. An end-to-end multi-task deep learning framework for skin lesion analysis. IEEE J Biomed Health Inform 2020; 24(10): 2912-21. http://dx.doi.org/10.1109/JBHI.2020.2973614 PMID: 32071016

[308] Jin Q, Cui H, Sun C, *et al.* Cascade knowledge diffusion network for skin lesion diagnosis and segmentation. Appl Soft Comput 2021; 99: 106881.

[309] Maron RC, Haggenmüller S, von Kalle C, *et al.* Robustness of convolutional neural networks in recognition of pigmented skin lesions. Eur J Cancer 2021; 145: 81-91. http://dx.doi.org/10.1016/j.ejca.2020.11.020 PMID: 33423009

[310] Wang X, Jiang X, Ding H, Zhao Y, Liu J. Knowledge-aware deep framework for collaborative skin lesion segmentation and melanoma recognition. Pattern Recognit 2021; 120: 108075. http://dx.doi.org/10.1016/j.patcog.2021.108075

[311] Liu L, Tsui YY, Mandal M. Skin lesion segmentation using deep learning with auxiliary task. J Imaging 2021; 7(4): 67. http://dx.doi.org/10.3390/jimaging7040067 PMID: 34460517

[312] Zhang J, Mei K, Zheng Y, Fan J. Learning multi-layer coarse-to-fine representations for large-scale image classification. Pattern Recognit 2019; 91: 175-89. http://dx.doi.org/10.1016/j.patcog.2019.02.024

[313] Coppola D, Lee HK, Guan C. Interpreting mechanisms of prediction for skin cancer diagnosis using multi-task learning. In: Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. 2020 Jun 14-19; Seattle, WA, USA. 734-5. http://dx.doi.org/10.1109/CVPRW50498.2020.00375

[314] Alzahrani S, Al-Nuaimy W, Al-Bander B. Seven-point checklist with convolutional neural networks for melanoma diagnosis. In: 2019 8th European Workshop on Visual Information Processing (EUVIP). 2019 Oct 28-31; Roma, Italy. 211-6.

[315] Kong Z, He M, Luo Q, *et al.* Multi-task classification and segmentation for explicable capsule endoscopy diagnostics. Front Mol Biosci 2021; 8: 614277. http://dx.doi.org/10.3389/fmolb.2021.614277 PMID: 34490342

[316] Chu T, Li X, Vo H V, *et al.* Improving weakly supervised lesion segmentation using multi-task learning. Medical Imaging with Deep Learning 2021.

[317] Jin C, Yu H, Ke J, *et al.* Predicting treatment response from longitudinal images using multi-task deep learning. Nat Commun 2021; 12(1): 1851. http://dx.doi.org/10.1038/s41467-021-22188-y PMID: 33767170

[318] LabelImg. Available from: https://github.com/tzutalin/labelImg/

[319] LabelMe. Available from: http://labelme.csail.mit.edu/Release3.0/

[320] Tokuoka Y, Suzuki S, Sugawara Y. An inductive transfer learning approach using cycle-consistent adversarial domain adaptation with application to brain tumor segmentation. In: Proceedings of the 2019 6th International Conference on Biomedical and Bioinformatics Engineering. 2019 Nov 13-15; Shanghai, China. 44-8. http://dx.doi.org/10.1145/3375923.3375948

[321] Dupre R, Fajtl J, Argyriou V, Remagnino P. Improving dataset volumes and model accuracy with semi-supervised iterative self-learning. IEEE Trans Image Process 2019; 29: 4337-48. http://dx.doi.org/10.1109/TIP.2019.2913986 PMID: 31059446

[322] Wei X, Wei X, Kong X, Lu S, Xing W, Lu W. FMixCutMatch for semi-supervised deep learning. Neural Netw 2021; 133: 166-76. http://dx.doi.org/10.1016/j.neunet.2020.10.018 PMID: 33217685

[323] Berthelot D, Carlini N, Goodfellow I, *et al.* Mixmatch: A holistic approach to semi-supervised learning. arXiv 2019.

[324] He Y, Shi J, Wang C, *et al.* Semi-supervised skin detection by network with mutual guidance. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019 Oct 27-Nov 2; Seoul, Korea (South). 2111-20. http://dx.doi.org/10.1109/ICCV.2019.00220

[325] Liu Y, Lee J, Park M, *et al.* Learning to propagate labels: Transductive propagation network for few-shot learning. arXiv 2018.

[326] Abubakar A, Ajuji M, Usman Yahya I. Comparison of deep transfer learning techniques in human skin burns discrimination. Appl Syst Innov 2020; 3(2): 20. http://dx.doi.org/10.3390/asi3020020

[327] Hosny KM, Kassem MA, Foaud MM. Skin melanoma classification using ROI and data augmentation with deep convolutional neural networks. Multimedia Tools Appl 2020; 79(33): 24029-55. http://dx.doi.org/10.1007/s11042-020-09067-2

[328] Hekler A, Kather JN, Krieghoff-Henning E, *et al.* Effects of label noise on deep learning-based skin cancer classification. Front Med (Lausanne) 2020; 7: 177.

http://dx.doi.org/10.3389/fmed.2020.00177 PMID: 32435646

[329] Zunair H, Ben Hamza A. Melanoma detection using adversarial training and deep transfer learning. Phys Med Biol 2020; 65(13): 135005.
http://dx.doi.org/10.1088/1361-6560/ab86d3 PMID: 32252036

[330] Marcus G, Davis E. Rebooting AI: Building artificial intelligence we can trust. New York City: Knopf Doubleday Publishing Group 2019.

[331] Li X, Xu Y, Xiang F, *et al.* Prediction of IDH mutation status of glioma based on multimodal MRI images. In: 2021 3rd International Conference on Intelligent Medicine and Image Processing. 2021 Apr 23-26; Tianjin, China. 39-44.

[332] Huang F, Zhang X, Zhao Z, Xu J, Li Z. Image–text sentiment analysis *via* deep multimodal attentive fusion. Knowl Base Syst 2019; 167: 26-37.
http://dx.doi.org/10.1016/j.knosys.2019.01.019

[333] Zadeh A, Chen M, Poria S, *et al.* Tensor fusion network for multimodal sentiment analysis. arXiv 2017.

[334] Li X, Xu Y, Xiang F, Liu Q, Huang W, Xie B. KINET: A non-invasive method for predicting ki67 index of glioma. In: 2021 IEEE International Conference on Image Processing (ICIP). 2021 Sep 19-22; Anchorage, AK, USA. 150-4.
http://dx.doi.org/10.1109/ICIP42928.2021.9506741

[335] Liu Z, Shen Y, Lakshminarasimhan VB, *et al.* Efficient low-rank multimodal fusion with modality-specific factors. arXiv 2018.

[336] Hou M, Tang J, Zhang J, *et al.* Deep multimodal multilinear fusion with high-order polynomial pooling. Adv Neural Inf Process Syst 2019; 32: 12136-45.

[337] Zadeh A, Liang PP, Mazumder N, *et al.* Memory fusion network for multi-view sequential learning. Proc Conf AAAI Artif Intell 2018; 32(1): 5634-41.

[338] Xu N, Mao W, Chen G. Multi-interactive memory network for aspect based multimodal sentiment analysis. Proc Conf AAAI Artif Intell 2019; 33(01): 371-8.
http://dx.doi.org/10.1609/aaai.v33i01.3301371

[339] Zhang Z, Chen K, Wang R, *et al.* Neural machine translation with universal visual representation. In: International Conference on Learning Representations. 2020 Apr 30; Addis Ababa, Ethiopia.

[340] Lu Y, Wu Y, Liu B, *et al.* Cross-modality person re-identification with shared-specific feature transfer. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2020 Jun 13-19; Seattle, WA, USA. 13379-89.
http://dx.doi.org/10.1109/CVPR42600.2020.01339

[341] Li X, Wang C, Tan J, *et al.* Adversarial multimodal representation learning for click-through rate prediction. In: Proceedings of The Web Conference 2020. 2020 Apr 20-24; Taipei, Taiwan. 827-36.
http://dx.doi.org/10.1145/3366423.3380163

[342] Qin Q, Hu W, Liu B. Feature projection for improved text classification. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020 Jul; 8161-71.
http://dx.doi.org/10.18653/v1/2020.acl-main.726

[343] Yang H, Wang T, Yin L. Adaptive Multimodal Fusion for Facial Action Units Recognition. Proceedings of the 28th ACM International Conference on Multimedia. 2020 Oct 12-16; Seattle, WA, USA. 2982-90.
http://dx.doi.org/10.1145/3394171.3413538

[344] Pérez-Rúa JM, Vielzeuf V, Pateux S, *et al.* Mfas: Multimodal fusion architecture search. In: Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019 Jun 15-19; Long Beach, CA, USA. 6966-75.

[345] Joze HRV, Shaban A, Iuzzolino ML, *et al.* MMTM: Multimodal transfer module for CNN fusion. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2020 Jun 13-19; Seattle, WA, USA. 13289-99.

[346] Fan X, Dai M, Liu C, *et al.* Effect of image noise on the classification of skin lesions using deep convolutional neural networks. Tsinghua Sci Technol 2019; 25(3): 425-34.
http://dx.doi.org/10.26599/TST.2019.9010029

[347] Hu L, Wang S, Li L, *et al.* How functions evolve in deep convolutional neural network. In: 2018 14th IEEE International Conference on Signal Processing (ICSP). 2018 Aug 12-16; Beijing, China. 1133-8.

[348] Chen CLP, Liu Z. Broad learning system: An effective and efficient incremental learning system without the need for deep architecture. IEEE Trans Neural Netw Learn Syst 2018; 29(1): 10-24.
http://dx.doi.org/10.1109/TNNLS.2017.2716952 PMID: 28742048

[349] Pintelas E, Liaskos M, Livieris IE, Kotsiantis S, Pintelas P. A novel explainable image classification framework: Case study on skin cancer and plant disease prediction. Neural Comput Appl 2021; 33(22): 1-19.
http://dx.doi.org/10.1007/s00521-021-06141-0

[350] Fan FL, Xiong J, Li M, Wang G. On interpretability of artificial neural networks: A survey. IEEE Trans Radiat Plasma Med Sci 2021; 5(6): 741-60.
http://dx.doi.org/10.1109/TRPMS.2021.3066428

[351] Stieler F, Rabe F, Bauer B. Towards domain-specific explainable AI: Model interpretation of a skin image classifier using a human approach. In: Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2021 Jun 19-25; Nashville, TN, USA. 1802-9.
http://dx.doi.org/10.1109/CVPRW53098.2021.00199

[352] Jiang S, Li H, Jin Z. A visually interpretable deep learning framework for histopathological Image-based skin cancer diagnosis. IEEE J Biomed Health Inform 2021; 25(5): 1483-94.
http://dx.doi.org/10.1109/JBHI.2021.3052044 PMID: 33449890

[353] Olah C, Mordvintsev A, Schubert L. Feature visualization: How neural networks build up their understanding of images. Distill 2017.

[354] Sanh V, Debut L, Chaumond J, *et al.* DistilBERT, a distilled version of BERT. arXiv 2019.

[355] Jiao X, Yin Y, Shang L, *et al.* Distilling bert for natural language understanding. arXiv 2019.

[356] Vandenhende S, Georgoulis S, Van Gansbeke W, Proesmans M, Dai D, Van Gool L. Multi-Task Learning for Dense Prediction Tasks: A Survey. IEEE Trans Pattern Anal Mach Intell 2021; PP: 1.
http://dx.doi.org/10.1109/TPAMI.2021.3054719 PMID: 33497328