

REVIEW

Open Access



# Opportunities and Challenges: Classification of Skin Disease Based on Deep Learning

Bin Zhang<sup>1,2</sup>, Xue Zhou<sup>1,2</sup>, Yichen Luo<sup>1,2</sup>, Hao Zhang<sup>1,2</sup>, Huayong Yang<sup>1,2</sup>, Jien Ma<sup>3</sup> and Liang Ma<sup>1,2\*</sup> 

## Abstract

Deep learning has become an extremely popular method in recent years, and can be a powerful tool in complex, prior-knowledge-required areas, especially in the field of biomedicine, which is now facing the problem of inadequate medical resources. The application of deep learning in disease diagnosis has become a new research topic in dermatology. This paper aims to provide a quick review of the classification of skin disease using deep learning to summarize the characteristics of skin lesions and the status of image technology. We study the characteristics of skin disease and review the research on skin disease classification using deep learning. We analyze these studies using datasets, data processing, classification models, and evaluation criteria. We summarize the development of this field, illustrate the key steps and influencing factors of dermatological diagnosis, and identify the challenges and opportunities at this stage. Our research confirms that a skin disease recognition method based on deep learning can be superior to professional dermatologists in specific scenarios and has broad research prospects.

**Keywords:** Skin disease, Image method, Deep learning, Disease classification

## 1 Introduction

Skin lesions are a common disease that cause suffering, some of which can have serious consequences, for millions of people globally [1]. Because of its complexity, diversity, and similarity, skin disease can only be diagnosed by dermatologists with long-term clinical experience and is rarely reproducible. It is likely to be misdiagnosed by an inexperienced dermatologist, which can exacerbate the condition and impede appropriate treatment. Thus, it is necessary to provide a quick and reliable method to assist patients and dermatologists in data processing and judgment.

Advances in deep learning have influenced numerous scientific and industrial fields and have realized significant achievements with inspiration from the human nervous system. With the rapid development of deep learning in biomedical data processing, numerous specialists have adopted this technique to acquire more

precise and accurate data. With the rapid increase in the amount of available biomedical data including images, medical records, and omics, deep learning has achieved considerable success in a number of medical image processing problems [2–4]. In this regard, deep learning is expected to influence the roles of image experts in biomedical diagnosis owing to its ability to perform quick and accurate assessments. This paper presents the characteristics of skin lesions, overviews image techniques, generalizes the developments in deep learning for skin disease classification, and discusses the limitations and direction of automatic diagnosis.

## 2 Features of Skin Disease

The skin is the largest organ of the human body; in adults, it can typically weigh 3.6 kg and cover 2 m<sup>2</sup> [5]. Skin guards the body against extremes of temperature, damaging sunlight, and harmful chemicals. As a highly organized structure, it consists of the epidermis, dermis, and hypodermis, providing the functions of protection, sensation, and thermoregulation [6]. The epidermis, the outermost layer of the skin, provides an excellent aegis

\*Correspondence: liangma@zju.edu.cn

<sup>1</sup> State Key Laboratory of Fluid Power and Mechatronic Systems, Zhejiang University, Hangzhou 310058, China

Full list of author information is available at the end of the article

to avoid environmental aggression. The dermis, beneath the epidermis, contains tough connective tissue, hair follicles, and sweat glands, which leads to the differentiation of skin appearance [7]. There are numerous causes of skin disease, including physical factors such as light, temperature, and friction, and biological factors such as insect bites, allergic diseases, and even viral infections. Environmental and genetic factors can also lead to the occurrence of skin diseases. In lesion imaging, complicating difficulties can include variations in skin tone, presence of artifacts such as hair, air bubbles, non-uniform lighting, and the physical location of the lesion. Moreover, the majority of lesions vary in terms of color, texture, shape, size, and location in an image frame [8]. There are 5.4 million new skin cancer patients in America every year. As of 2014, there were 420 million people globally suffering from skin disease, including nearly 150 million people in China, the population of which accounts for 22% of the world's population, yet medical resources account for only 2%. Influenced by the living environment, areas with reduced economic development and poverty are more prone to skin disease. The high cost of treatment, repeated illness occurrences, and delays in treatment

have focused attention on the requirement for healthy survival and social development. The high cost of treatment, repeated illness occurrences, and delays in treatment have brought challenges to the healthy survival and social development.

The accurate diagnosis of a particular skin disease can be a challenging task, mainly for the following reasons. First, there are numerous kinds of dermatoses, nearly 3000 recorded in the literature. Stanford University has developed an algorithm to demonstrate generalizable classification with a new dermatologist-labeled dataset of 129450 clinical images divided into 2032 categories [9]. Figure 1 displays a subset of the full taxonomy; this has been organized clinically and visually by medical experts. Secondly, the complex manifestation of the disease is also a major challenge for doctors. Morphological differences in the appearance of skin lesions directly influence the diagnosis mainly as there can be relatively poor contrast between different skin diseases, which cannot be distinguished without considerable experience. Finally, for different skin diseases, the lesions can be overly similar to be distinguished using only visual information. Different diseases can have similar manifestations and the same

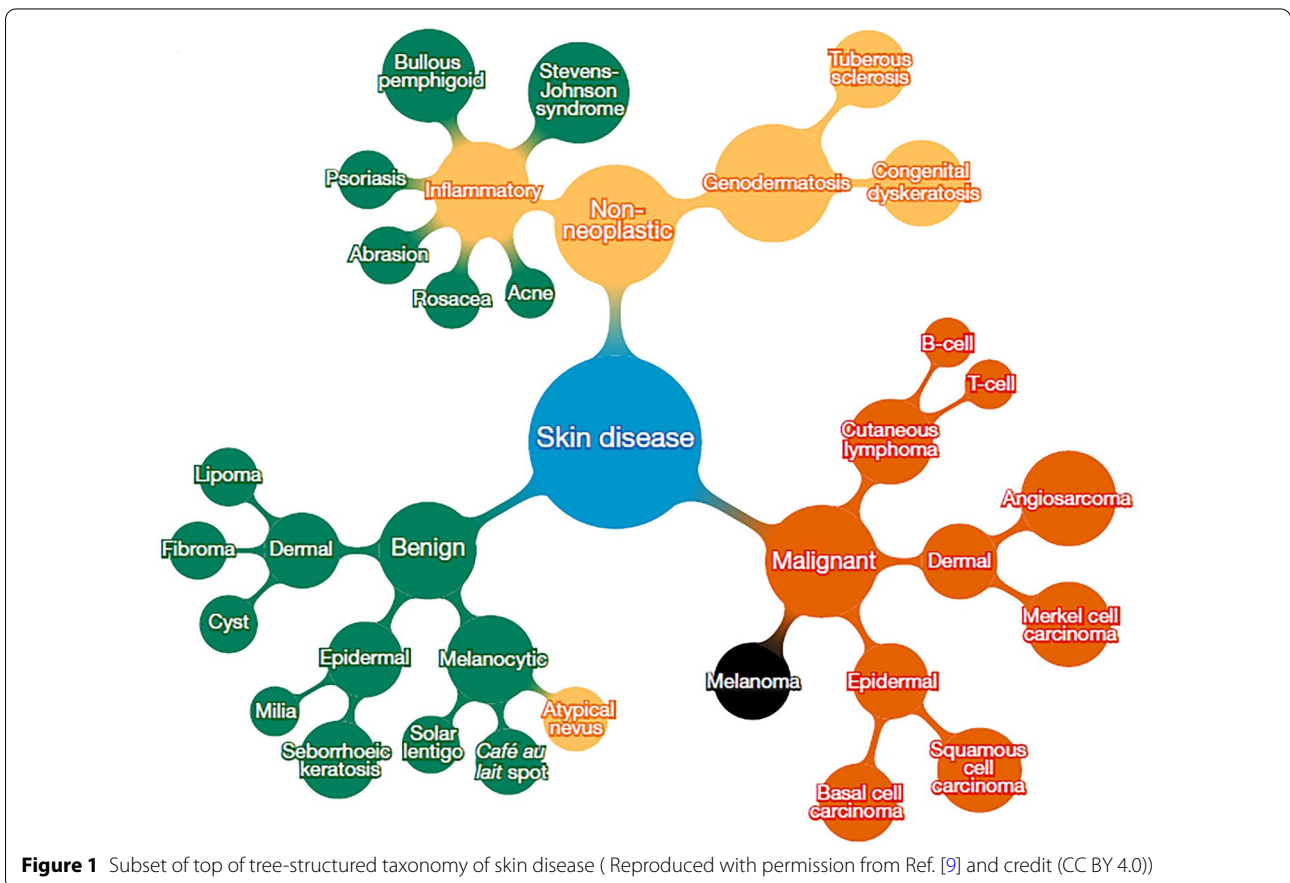


Figure 1 Subset of top of tree-structured taxonomy of skin disease (Reproduced with permission from Ref. [9] and credit (CC BY 4.0))

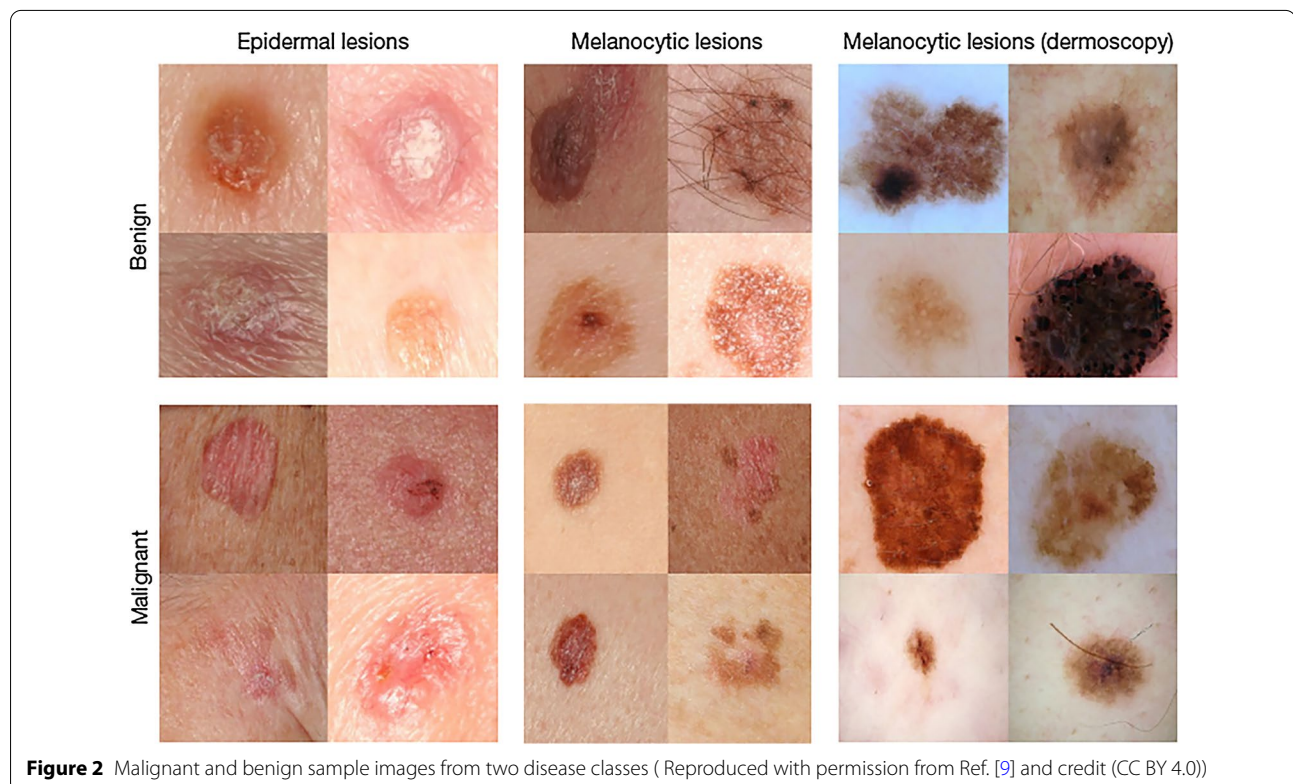
disease can have different manifestations in different people, body parts, and disease periods [10]. Figure 2 displays sample images demonstrating the difficulty in distinguishing between malignant and benign lesions, which share several visual features. Unlike benign skin diseases, malignant diseases, if not treated promptly, can lead serious consequences. Melanoma [11], for example, is one of the major and most fatal skin cancers. The five-year survival rate of melanoma can be greater than 98% if found in time; this figure in those where spread has occurred demonstrates a significant drop to 17% [12]. In 2015, there were 3.1 million active cases, representing approximately 70% of skin cancer deaths worldwide [13, 14].

The diagnosis of skin disease relies on clinical experience and visual perception. However, human visual diagnosis is subjective and lacks accuracy and repeatability, which is not found in computerized skin-image analysis systems. The use of these systems enables inexperienced operators to prescreen patients [15]. Compared with other diseases or applications such as industrial fault diagnosis, the visual manifestation of skin disease is more prominent, facilitating the significant value of deep learning in image recognition with visual sensitivity. Through the study of large detailed images, dermatology can become one of the most suitable medical fields for telemedicine and artificial intelligence (AI). Using imaging

methods, it could be possible for deep learning to assist or even replace dermatologists in the diagnosis of skin disease in the near future.

### 3 Image Methods

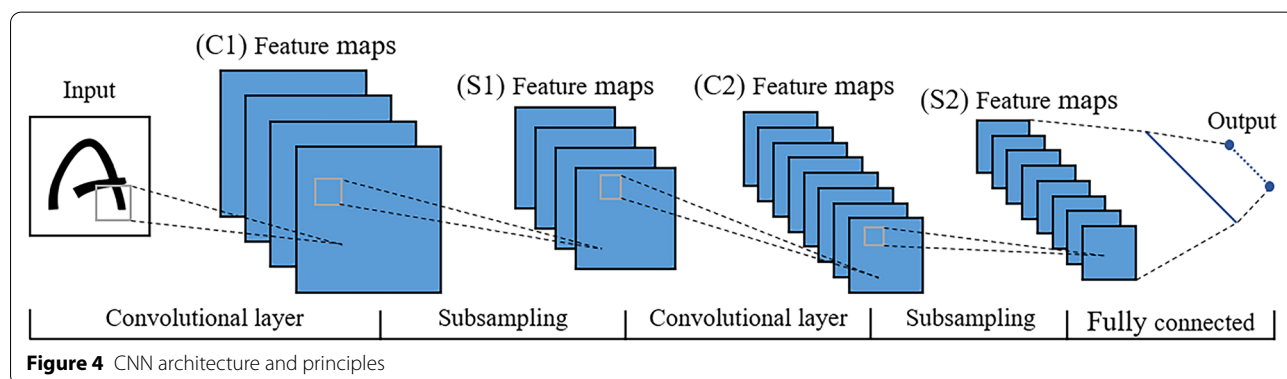
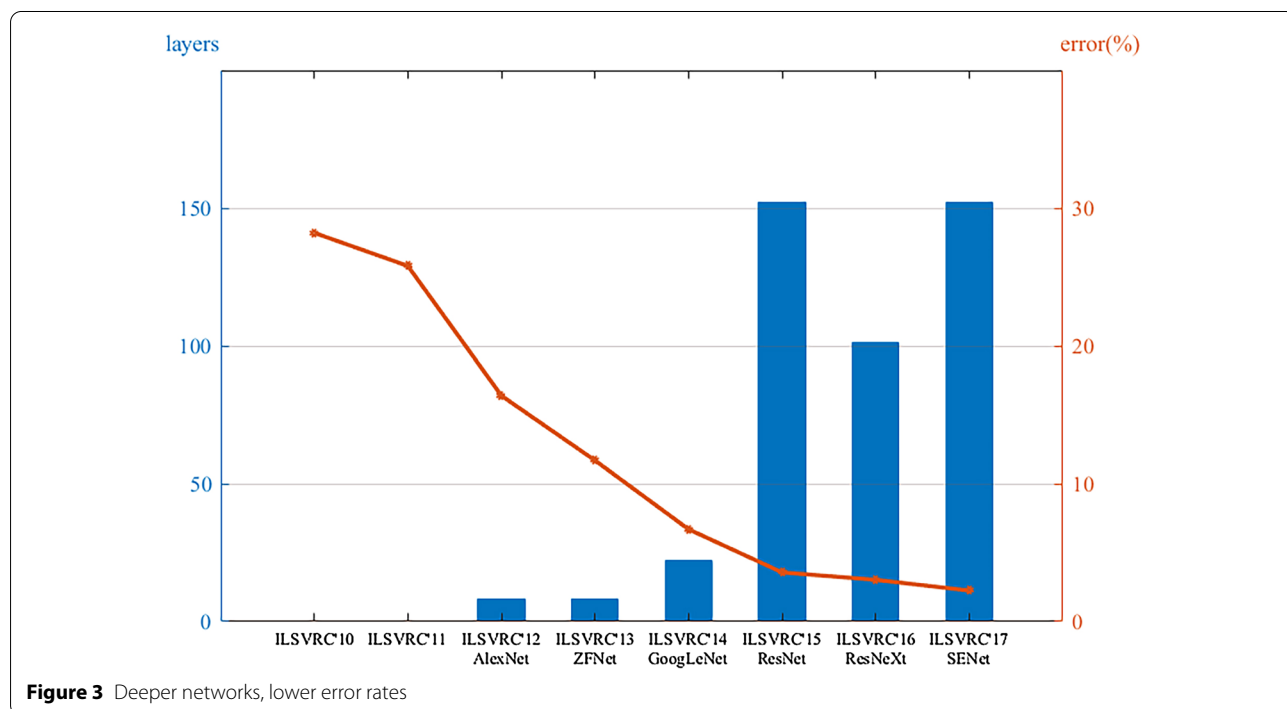
Deep learning is a class of machine learning that automatically learns hierarchical features of data using multiple layers composed of simple and nonlinear modules. It transforms the data into representations that are important for discriminating the data [16]. As early as 1998, the LeNet network was proposed for handwritten digital recognition [17]. However, owing to the lack of computational power, it was difficult to support the required computation. Until 2012, this method was successfully applied and overwhelmingly outperformed previous machine learning methods for visual recognition tasks at a competitive challenge in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [18, 19]. This was a breakthrough that used convolutional networks to virtually halve the error rate for object recognition, and precipitated the rapid adoption of deep learning by the computer vision community [16]. Since then, deep learning algorithms have undergone considerable development because of the improved capabilities of hardware such as graphics processing units (GPUs). Different models, such as ZFNet [20], VGG [21], GoogLeNet [22], and



ResNet [23], have been proposed. The top-5 error rate in ImageNet dropped from 16.4% in 2012 to 2.25% in 2017 (Figure 3); correspondingly, that of humans was approximately 5%. It has dramatically improved tasks in different scientific and industrial fields including not only computer vision but also speech recognition, drug discovery, clinical surgery, and bioinformatics [24–26].

The structure of a convolutional neural network (CNN), which is a representative deep learning algorithm, is displayed in Figure 4. The actual model is similar to this figure, in addition to deeper layers and more convolution kernels. A CNN is a type of “feedforward neural network” inspired by human visual perception mechanisms, and can learn a large number of mappings between inputs and outputs without any precise mathematical expression between them. The first convolutional

filter of the CNN is used to detect low-order features such as edges, angles, and curves. As the convolutional layer increases, the detected features become more complex [20]. The pooling layer, or named subsampling layer, converts a window into a pixel by taking the maximum or average value [27], which can reduce the size of the feature map. After the image passes the last fully connected layer, the model maps the learned distributed feature to the sample mark space and provides the final classification type. The layout of the CNN is similar to the biological neural network, with sparse structures and shared weights, which can reduce the number of parameters and improve the fitting effect to prevent overfitting. Deep CNNs demonstrate the potential for variable tasks across numerous fine-grained object categories and have unique advantages in the field of image recognition.





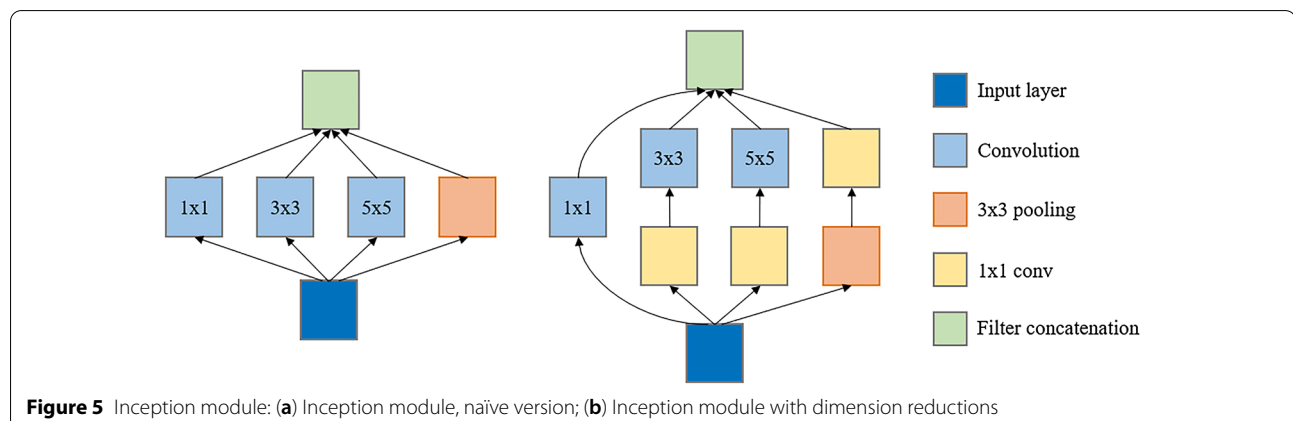
The selection of a suitable model is crucial. The GoogLeNet model, with a structure called inception (Figure 5), is proposed which can not only maintain the sparsity of the network structure but can also use the high computational performance of the dense matrix [22]. GoogLeNet has been learned and used by numerous researchers because of its excellent performance. Therefore, the Google team has further explored and improved it, resulting in an upgraded version of GoogLeNet, Inception v3 [28], which has become the first choice for current research. With Google's Inception v3 CNN architecture pretrained to a high-level accuracy on the 1000 object class of ImageNet, researchers can remove the final classification layer from the network, retrain it with their own dataset, and fine-tune the parameters across all the layers.

Google's TensorFlow [29], Caffe [30], and Theano [31] deep learning frameworks can be used for training. Theano is a Python library and optimizing compiler for manipulating and evaluating mathematical expressions. It pioneered the trend of using symbolic graphs for programming a network; however, it lacks a low-level interface and the inefficiency of the Python interpreter limits its usage. Caffe's ConvNet implementation with numerous extensions being actively added is excellent; however, its support for recurrent networks and language modeling in general is poor. If both CPU and GPU supports are required, additional functions must be implemented. Specifying a new network is fairly easy in TensorFlow using a symbolic graph of vector operations; however, it has a major weakness in terms of modeling flexibility. It has a clean, modular architecture with multiple frontends and execution platforms, and the library can be compiled on Advanced RISC Machines (ARM).

Deep learning has been gradually applied to medical image data, as medical image analysis approaches are considerably similar to computer vision techniques

[32]. Although numerous studies were initially undertaken using relatively small datasets and a pretrained deep learning model as a feasibility study, a robust validation of the medical application is required [33–35]. Hence, big data from medical images have been collected to validate the feasibility of medical applications [9, 36]. For example, Google researchers collected large datasets consisting of more than 120,000 retinal fundus images for diagnosing diabetic retinopathy and demonstrated high sensitivity and specificity for detection [37].

Owing to the development of hardware and advancement of the algorithms, deep learning now includes considerably more functionality than could previously be imagined. Researchers are now more likely to predict and distinguish what is difficult to diagnose with complex mechanisms and similar characterizations [38, 39]. Deep learning is a powerful machine learning algorithm for classification while extracting low- to high-level features [40, 41]. A key difference in deep learning compared to other diagnostic methods is its self-learning nature. The neural network is not designed by humans; rather, it is designed by the data itself. Table 1 presents several published achievements on disease diagnosis using pictures or clinical images, which proves that deep learning can be compared with professional specialists in certain fields. Furthermore, many researchers have indicated interest in mobile diagnostics that allow the use of mobile technology. Smartphones with sufficient computing power and fast development to extend the versatility and utility could be used to scan, calculate, analyze anytime and anywhere to detect skin disease [42–44]. Researchers have developed such a system based on AI that allows users to install apps on their smartphones and analyze and judge suspicious lesions on the body by taking a picture [45].



**Table 1** Selected applications of deep learning

Disease type	Diagnostic target	Architecture	Accuracy	Control group
Skin cancer [9]	Multiple classification	Inception v3 CNN	> 0.960	~ 0.880
Congenital cataract [35]	Binary classification, object localization	VGG	0.976–0.989	–
Pneumonia [36]	Multiple classification, object detection	CheXNet	0.735–0.937	0.633–0.914
Diabetic retinopathy [37]	Binary classification	Inception v3	0.840	0.800–0.910
Cardiovascular disease [38]	Binary classification	–	0.764	0.728
Alzheimer's disease [40]	Binary classification	LeNet, GoogLeNet	0.845–0.988	–
Genetic disorders [46]	Binary classification	DeepGestalt	0.920–0.969	0.710–0.870
Cutaneous tumors[47]	Multiple classification	ResNet-152	0.96	–
Melanoma[48]	Binary classification	Inception v4	0.860	0.790
Malaria parasites [49, 50]	Binary classification, logistic regression	CNN	0.897	–
Cardiac contractile dysfunction [51]	Binary classification	CNN	0.835–0.925	–
Arrhythmia [52]	Multiple classification	DNN	0.837	0.780
Melanoma and Seborrheic keratosis [53]	Binary classification	DCNN	0.917	0.886
Skin tumor [54]	Multiple classification	DCNN	0.924	0.853
Melanoma [55]	Binary classification	ResNet-152	0.944	0.823

#### 4 Skin Disease Classification Using Deep Learning

Using the deep learning technique, the pattern recognition of images can be performed automatically once the program is established. Images can be input to a CNN with high fidelity and important features can be automatically obtained. Therefore, information extraction from images prior to the learning process is not necessary with this technique. In shallow layers, simple features such as the edges within the images are learned. At deep layers near the output layer, more complex high-order features are learned [56]. Different researchers, institutions, and challenges are working on the automatic diagnosis of skin disease, and different deep learning methods have been developed for the recognition of dermatological disease; these have been proven to be effective in numerous fields [57]. For example, the International Skin Imaging Collaboration (ISIC) is a challenge that focuses on the automatic analysis of skin lesions. The goal of the challenge (started in 2017) is to support the research and development of algorithms for the automated diagnosis of melanoma including lesion segmentation, dermoscopic feature detection within a lesion, and classification of melanoma [58, 59], which is also the main goal in the field of dermatology [60]. In general, this method is a modeling framework that can learn the functional mapping from the input images to output. The input image is a preprocessed image; the output image is a segmentation mask. The network structure involves a series of convolution and pooling layers, followed by a fully connected

layer, followed by a series of unpooling and disconnection operations [61].

The diagnosis of skin diseases typically consists of four components: image acquisition, image preprocessing, feature extraction and classification, and evaluation of the criteria. Image acquisition is the basis for skin classification, and more images typically indicate greater accuracy and better adaptability (for the data size of selected projects, please refer to Table 2). Preprocessing is used to crop and zoom the images and segment lesions for better training. Feature extraction mainly acquires the features of the skin lesions through color, texture, and boundary information. The evaluation of the results is the final step, which is used to judge whether the classification model is reasonable and achieves its objective.

##### 4.1 Image Acquisition

Deep learning requires a large number of images to extract disease features. These datasets are typically available from the Internet, open dermatology databases, and hospitals in collaboration with research units, and are labeled by professional dermatologists after removing blurry and distant images. An excellent dataset should be composed of dermoscopic images. Dermoscopy is a non-invasive skin imaging technology that can observe the skin structure at the junction of the epidermis and dermis, and clearly indicate the nature, distribution, arrangement, edge, and shape of pigmented skin lesions. Because of the uncertainty of imaging conditions, such as shooting angle, illumination, and storage pixels, the imaging

**Table 2** Data size of selected deep learning projects

Disease type	Images type	Database	Training set	Testing set
Skin cancer [9]	Clinical and dermoscopy images	129450	127463	1942
Cutaneous tumors [47]	Clinical images	17125	15408	1276
Genetic disorders [46]	Facial images	17000	1693	32
Melanoma [62]	Dermoscopic images	1250	900	350
Melanoma [53]	Dermoscopic images	2750	2000	600
Skin tumor [54]	Clinical images	6009	4867	1142

**Table 3** Skin disease datasets

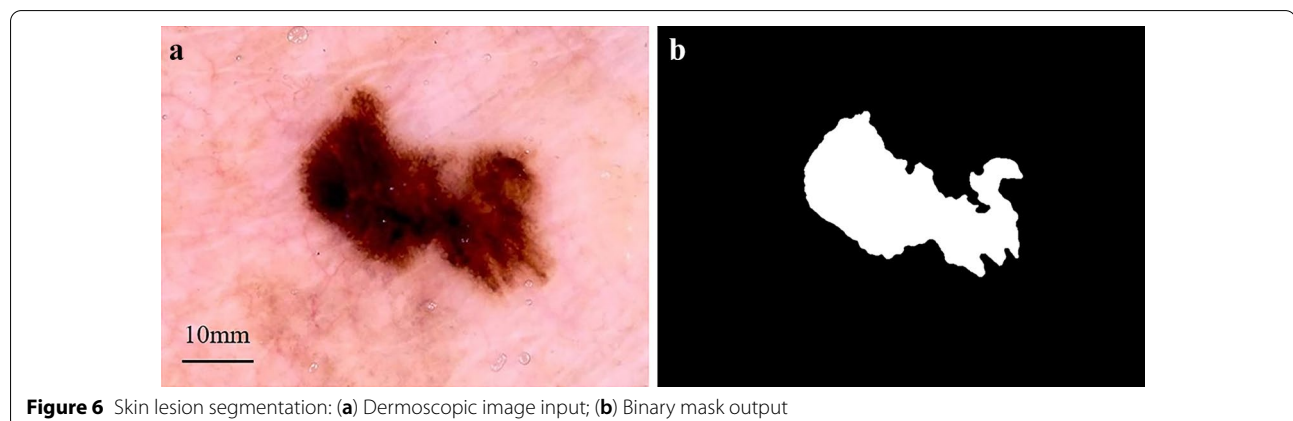
Name	Number of images	Disease classification	Image type
BCN2000 [64]	19424	8	Dermoscopic images
HAM10000 [65]	10015	7	Dermoscopic images
ISIC Archive [66]	23906	7	Dermoscopic images
ISIC 2016 [58]	1279	2	Dermoscopic images
ISIC 2017 [67]	2000	3	Dermoscopic images
ASan [47]	17125	12	Clinical images
Pigmentary dermatosis [68]	12816	6	Clinical images
DermQuest [69]	–	–	Clinical images
Dermnet [70]	–	–	Clinical images

effect of non-dermoscopic images can be influenced. Selected published datasets are listed in Table 3 covering more than a dozen kind of skin diseases, among which melanoma has the greatest probability of occurrence. However, owing to the lack of a unified standard for skin disease images, the labeling of images is time-consuming and labor-intensive, which significantly limits the size of the current public datasets. Therefore, numerous studies have combined multiple datasets for use [43, 63].

## 4.2 Image Preprocessing

Effective image quality can improve the generalization ability of a model. Preprocessing can reduce irrelevant information in the image, improve the intensity of the relevant information, simplify the data, and improve the reliability. The general image preprocessing process is as follows:

- (1) Image segmentation. Skin lesion segmentation is the essential step for the majority of classification tasks. Accurate segmentation contributes to the accuracy, computation time, and error rate of subsequent lesion classification [71, 72]. It is crucial for image analysis for the following two reasons. First, the border of a lesion provides important information for accurate diagnosis, including numerous clinical features such as asymmetry and border irregularity. Secondly, the extraction of other important clinical features such as atypical dots and color variegation critically depends on the accuracy of the border detection [8, 73]. Given a inputted dermoscopic image (Figure 6a), the goal of the segmentation process is to generate a two-dimensional mask (Figure 6b) that provides an accurate separation between the lesion area and surrounding healthy skin [74].

**Figure 6** Skin lesion segmentation: (a) Dermoscopic image input; (b) Binary mask output

- (2) **Resize.** Lesions frequently occupy a relatively small area, although skin images can be considerably large [75, 76]. Before this task, images for a deep learning network should be preprocessed because the resolution of the original lesion images is typically overly large, which entails a high computation cost [77]. Accurate skin lesion segmentation enhances its capability by incorporating a multiscale contextual information integration scheme [62]. To avoid distorting the shape of the skin lesion, the images should be cropped to the center area first and then proportionally resized. Images are frequently resized to  $224 \times 224$  or  $227 \times 227$  pixels through scaling and clipping [78], which is the appropriate size after combining the amount of calculation and information density.
- (3) **Normalization.** The image data are mapped to the interval of  $[0,1]$  or  $[-1,1]$  in the same dimension. The essence of normalization is a kind of linear transformation that does not cause “failure” after changing the data. Conversely, it can improve the performance of the data, accelerate the solution speed of gradient descent, and enhance the convergence speed of the model.
- (4) **Data augmentation.** Owing to privacy and professional equipment problems, it is difficult to collect sufficient data in the process of skin disease identification. A data set that is overly small can easily lead to overfitting owing to the lack of learning ability of the model, which makes the network model lack generalization ability. A method called data augmentation is adopted to expand the dataset to meet the requirements of deep learning for big data, such as rotation, random cropping, and noise [79]. Figure 7 displays several methods of image processing by which the image database can be extended to meet the training requirements.

### 4.3 Feature Extraction and Classification

Early detection of lesions is a crucial step in the field of skin cancer treatment. There is a significant benefit if this can be achieved without penetrating the body. Feature extraction of skin disease is an important tool that can be used to properly analyze and explore an image [80]. Feature extraction can be simply viewed as a dimensionality reduction process; that is, converting picture data into a vector of a certain dimension with picture features. Before deep learning, this was typically determined manually by dermatologists or researchers after investigating a large number of digital skin lesion images. A well-known method for feature extraction is based on the

ABCD rule of dermoscopy. ABCD stands for asymmetry, border structure, color variation, and lesion diameter. It defines the basis for disease diagnosis [81]. The extracted and fused traits such as color, texture, and Histogram of Oriented Gradient (HOG) are applied subsequently with a serial-based method. The fused features are selected afterwards by implementing a novel Boltzman entropy method [82], which can be used for the early detection. However, this typically has enormous randomness and depends on the quantity and quality of the pictures, as well as the experience of the dermatologists.

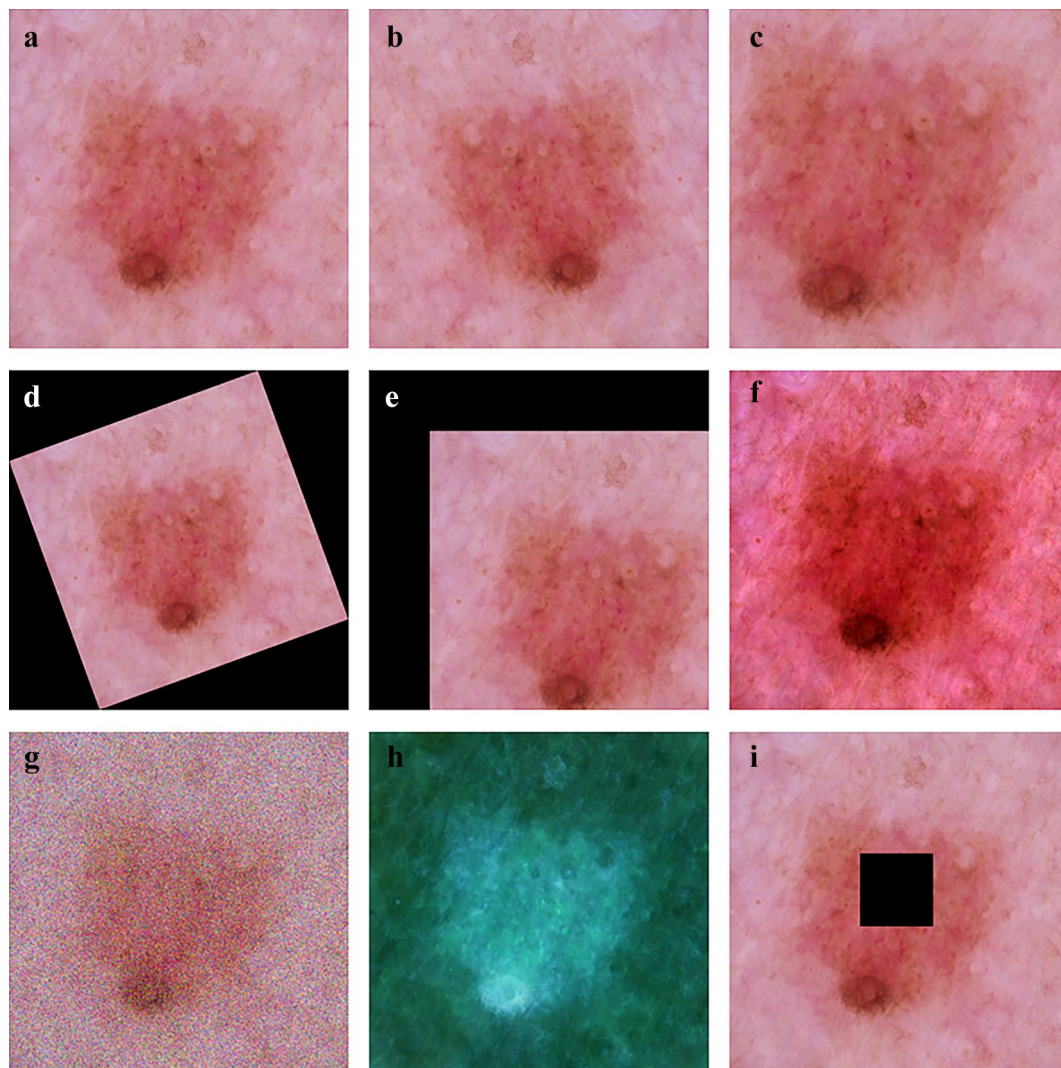
From a classification perspective, feature extraction has numerous benefits: (i) reducing classifier complexity for better generalization, (ii) improving prediction accuracy, (iii) reducing training and testing time, and (iv) enhancing the understanding and visualization of the data. The mechanism of neural networks is considerably different from that of traditional methods. Visualization indicates that the first layers are essentially calculating edge gradients and other simple operations such as SIFT [83] and HOG [84]. The folded layers combine the local patterns into a more global pattern, ultimately resulting in a more powerful feature extractor. In a study using nearly 130000 clinical dermatology images, 21 certified dermatologists tested the skin lesion classification with a single CNN, directly using pixels and image labels for end-to-end training; this had an accuracy of 0.96 for carcinoma [9]. Subsequently, researchers used deep learning to develop an automated classification system for 12 skin disorders by learning the abnormal characteristics of a malignancy and determined visual explanations from the deep network [47]. A third study combined deep learning with traditional methods such as hand-coded feature extraction and sparse coding to create a collection for melanoma detection that could yield higher performance than expert dermatologists. These results and others [85–87] confirm that deep learning has significant potential to reduce doctors’ repetitive work. Despite problems, it would be a significant advance if AI could reliably simulate experienced dermatologists.

### 4.4 Evaluation Criteria and Benchmarking

Evaluation and criterion, typically based on the following three points, reliability, time consumption, and training and validation are vital in this field [88]. Researchers [73, 89, 90] have used all three criteria to develop and design methods and techniques for detecting and diagnosing skin disease. Others [71, 91, 92] have used only two criteria, reliability, and training and validation to evaluate and discuss the different types of classifiers.

Numerous studies have demonstrated that acceptable reliability, time complexity, and error rates within a dataset cannot be achieved at the same time; hence,





**Figure 7** Data argumentation: (a) Original image; (b) Flip; (c) Random crop; (d) Rotation; (e) Shift; (f) Color jittering; (g) Noise; (h) Standardization; and (i) Paste

researchers must establish different standards. Once one of them is selected, the performance of the others diminishes [90, 93]. Consequently, conflicts among dermatological evaluation criteria pose a serious challenge to dermatological classification methods. These requirements must be considered during the evaluation and benchmarking. The dermatological classification method should standardize the requirements and objectives and use a programmatic process in research, evaluation, and benchmarking. Moreover, new flexible evaluations should address all conflicting standards and issues [94].

Despite the conflicts, important criteria are the key goals for evaluation and benchmarking. It is necessary to develop appropriate procedures for these goals while

increasing the importance of specific evaluation criteria and decreasing other standards [95]. When evaluating the results obtained using the diagnostic model, researchers must consider the quality of the dataset used to build the model and choose the parameters that can adjust that model. The time complexity and error rate in the dataset have proven to be important in the field of dermatology, which, with more consideration during the evaluation process, can optimize the consistency of the results [63]. In general, the goal is to obtain a balanced classifier for sensitivity and specificity.

## 5 Limitations and Prospect

### 5.1 Limitations

In general, the advantage of AI is that it can help doctors perform tedious repetitive tasks. For example, if sufficient blood is scanned, an AI-powered microscope can detect low-density infections in micrographs of standard, field-prepared thick blood films, which is considered to be time-consuming, difficult, and tedious owing to the low density and small parasite size and abundance of similar non-parasite objects [49]. The requirement for staff training and purchase of expensive equipment for creating dermoscopic images can be replaced by software using CNNs [96]. In the future, the clinical application of deep learning for the diagnosis of other diseases can be investigated. Transfer learning could be useful in developing CNN models for relatively rare diseases. Models could also evolve such that they require fewer preprocessing steps. In addition to these topics, a deeper understanding of the reconstruction kernel or image thickness could lead to improved deep learning model performance. Positive effects should continue to grow owing to the emergence of higher precision scanners and image reconstruction techniques [56]. However, we must realize that although AI has the ability to defeat humans in several specific fields, in general the performance of AI is considerably less than acceptable in the majority of cases [97]. The main reasons for this are as follows.

- (1) Medicine is an area that is not yet fully understood. Information is not completely transparent. The characteristics of dermatology determine that the majority of the data cannot be obtained. At the same time, the AI technology route is immature, the identification accuracy of which must be improved owing to the uncertainty of manual diagnosis. There is no strict correspondence between the symptoms and results of a disease and no clear boundary between the different diseases. Thus, the use of deep learning for disease diagnosis continues to require considerable effort.
- (2) Before systematic debugging, extensive simulation, and robust validation, flawed algorithms could harm patients, which could lead to medical ethical issues, and therefore require forward-looking scrutiny and stricter regulation [98]. As a “black box”, the principle of deep learning is unexplained at this stage, which could result in unpredictable system output. Moreover, it is possible that humans could not truly understand how a machine functions, even though it is actually inspired by humans [99, 100]. Hence, whether or not patient care can be accepted using an opaque algorithm remains a point of discussion.
- (3) There is a problem with the change in the error rate value in a dataset, which is caused by the change in the size of the dataset used in different skin cancer experiments. Therefore, the lack of a standard dataset can lead to serious problems; the error rate values are considered in many experiments. In addition, the collection of datasets for numerous studies depends on individual research, leading to unnecessary effort and time. When the actual class is manually marked and compared to the predicted class to calculate one of the parameter matrices, pixels are lost when the background is cut from the skin cancer image using Adobe Photoshop [101]. At this point, the process influences the results of all the parameter reliability groups (matrices, relationships, and behaviors), which are considered controversial. High reliability and low rate of time complexity cannot be achieved simultaneously, which is reflected in the training process and is influenced by conflicts between different standards, leading to considerable challenges [93, 102]. A method that works for the detection of one skin lesion could possibly not work for the detection of others [103]. Numerous different training and test sets have been used to evaluate the proposed methods. Moreover, for the parameters in the training and evaluation, different researchers are interested in different parts. This lack of uniformity and standardization across all papers makes a fair comparison virtually impossible [50, 104]. Although these indicators in the literature have been widely criticized, studies continue to use them to evaluate the application to skin cancer and other image processing fields.
- (4) The data used for evaluation are frequently overly small to allow a convincing statement regarding a system’s performance to be made. Although it is not impossible to collect an abundance of relevant data through the Internet in this information age, this information, with significant uncertainty, apparently cannot meet the requirements of independent and identical distribution, which is one of the important prerequisites for deep learning to be successfully applied. For certain rare diseases and minorities, only a limited number of images are available for training. To date, a large number of algorithms have demonstrated prejudice against minority groups, which could cause a greater gap in health service between the “haves” and the “have-nots” [105]. Numerous cases are required for the training process using deep learning techniques. In addition, although the deep learning technique has been successfully applied to other tasks, the developed models in skin are valid in only specific dedi-

cated diseases and are not applicable to common situations. Diagnosing dermatology is a complex process that, in addition to image recognition, must be supplemented by other means such as palpation, smell, temperature change, and microscopy.

## 5.2 Prospect

Deep learning has made considerable progress in the field of skin disease recognition. More attempts and explorations in the future can be considered in the following aspects.

### (1) Establishment of standardized skin disease image dataset

A large amount of data is the basis of skin disease recognition and the premise of acceptable generalization ability of the network model. However, the number of images, disease types, image size, and shooting and processing methods of the published datasets are considerably different, which leads to the confusion of different studies and the loss of the ability to quantitatively describe different models. Moreover, it is difficult to collect images of certain rare diseases. As mentioned above, there are numerous kinds of skin diseases; however, only approximately 20 datasets are available, including less than 20 kinds of skin diseases. There is an urgent requirement to expand access to medical images. For example, Indian researchers have trained neural networks to analyze images from “handheld imaging devices” instead of stationary dermatoscope devices to provide more prospects for early and correct diagnosis [106]. However, a public database that allows the collection of a sufficient number of labeled datasets is likely necessary to truly represent projections of the population.

### (2) Interpretability of skin disease recognition

The progress of deep learning in skin disease recognition depends on a highly nonlinear model and parameter adjustment technology. However, the majority of the neural networks are “black box” models, and their internal decision-making process is difficult to understand. This “end-to-end” decision-making mode leads to the weak explanatory power of deep learning. The internal logic of deep learning is not clear, which makes the diagnosis results of the model less convincing. The interpretability research of skin disease classification could allow the owner of the system to clearly know the behavior and boundary of the system, and ensure the reliability and safety of the system. Moreover, it could monitor the moral problems and viola-

tions caused by training data deviation and provide a superior mechanism to follow the requirements within an organization to solve the bias and audit problems caused by AI [107].

### (3) Intelligent diagnosis and treatment of skin diseases

Deep learning can be used to address the increasing number of patients with skin disease and relieve the pressure of limited dermatologists. With the popularity of mobile phones, mobile computers, and wearable devices, a skin disease recognition system based on deep learning can be expected to be available to intelligent devices to serve more people. Using a mobile device camera, users can upload their own photos of the affected area to the cloud recognition system and download the diagnosis results at any time. Through simple communication with the “skin manager”, diagnosis suggestions and possible treatment methods could be available. Furthermore, the “skin manager” could monitor the user’s skin condition and provide real-time protection methods and treatment suggestions.

Computer diagnostic systems can assist trained dermatologists rather than replace them. These systems can also be useful for untrained general practitioners or telemedicine clinics. For health systems, improving workflows could increase efficiency and reduce medical errors. Hospitals could make use of large-scale data and recommend data sharing with a cloud-based platform, thus facilitating multihospital collaboration [108]. For patients, it should be possible to enjoy the medical resources of the top hospitals in big cities in remote and less modernized areas by telemedicine or enabling them to process their own data [109].

## 6 Conclusions

The potential benefits of deep learning solutions for skin disease are tremendous and there is an unparalleled advantage in reducing the repetitive work of dermatologists and pressure on medical resources. Accurate detection is a tedious task that inevitably increases the demand for a reliable automated detection process that can be adopted routinely in the diagnostic process by expert and non-expert clinicians. Deep learning is a comprehensive subject that requires a wide range of knowledge in engineering, information, computer science, and medicine. With the continuous development of the above fields, deep learning is undergoing rapid development and has attracted the attention of numerous countries. Powered by more affordable solutions, software that can quickly collect and meaningfully process massive data, and hardware that can accomplish what people cannot, it is evident that deep learning for



the identification of skin disease is a potential technique in the foreseeable future.

#### Acknowledgements

Not applicable.

#### Authors' contributions

BZ was in charge of the whole trial; XZ wrote the manuscript; YCL assisted with structure and language of the manuscript. HZ designed the experiments; HYY assisted with experimental setup; LM and JEM jointly supervised this work; All authors read and approved the final manuscript.

#### Authors' information

Bin Zhang received his PhD degree in Engineering from Zhejiang University, China, in 2009. Currently, he is a researcher at School of Mechanical Engineering, Zhejiang University, China. His research interests include intelligent digital hydraulics and biological manufacturing based on fluid extrusion.

Xue Zhou received his B.S. degree in mechanical design and manufacturing from Xiamen University, China, in 2017. He is currently a PhD candidate at State Key Laboratory of Fluid Power Transmission and Control, Zhejiang University, China. His research mainly focuses on in vivo 3D printing and skin wound detection and repair.

Yichen Luo received his B.S. degree from Southeast University, China, in 2015. He is currently a PhD candidate at State Key Laboratory of Fluid Power Transmission and Control, Zhejiang University, China. His research mainly interests include 3D bioprinting and biological manufacturing.

Hao Zhang received his master degree from Zhejiang University, China, in 2019. Her research interests include deep learning and skin disease classification.

Huayong Yang received his PhD degree from the University of Bath, United Kingdom, in 1988. He is the dean of School of Mechanical Engineering, Zhejiang University, China. He was elected as an academicien of the Chinese Academy of Engineering in 2013. His research mainly focuses on energy saving of fluid power and electromechanical systems and 3D printing of biological organs.

Jien Ma received her PhD degree from Zhejiang University, China, in 2009. She is currently a professor at College of Electrical Engineering, Zhejiang University, China. Her research interests include electrician theory and new energy technology.

Liang Ma received his PhD degree from University of Washington, USA, 2012. He is currently as an associate professor at School of Mechanical Engineering, Zhejiang University, China. His research interests include organ bioprinting and Microfluidic chip.

#### Funding

Supported by Key Research and Development Projects of Zhejiang Province of China (Grant No. 2017C01054), National Key Research and Development Program of China (Grant No. 2018YFA0703000), National Natural Science Foundation of China (Grant No. 51875518), and Fundamental Research Funds for the Central Universities of China (Grant Nos. 2019XZZX003-02, 2019FZA4002).

#### Competing interests

The authors declare there is no competing interests.

#### Author Details

<sup>1</sup>State Key Laboratory of Fluid Power and Mechatronic Systems, Zhejiang University, Hangzhou 310058, China. <sup>2</sup>School of Mechanical Engineering, Zhejiang University, Hangzhou 310058, China. <sup>3</sup>College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China.

Received: 10 November 2020 Revised: 16 May 2021 Accepted: 8 October 2021

Published online: 24 November 2021

#### References

- [1] T Tarver. Cancer Facts & Figures 2012. American Cancer Society (ACS). *Journal of Consumer Health on the Internet*, 2012, 16(3): 366-367.
- [2] G M Weber, K D Mandl, I S Kohane. Finding the missing link for big biomedical data. *Jama*, 2014, 311(24): 2479.
- [3] D-M Filimon, A Albu. Skin diseases diagnosis using artificial neural networks. *2014 IEEE 9th IEEE International Symposium on Applied Computational Intelligence and Informatics (SACI)*, IEEE, 2014: 189-194, <https://doi.org/10.1109/SACI.2014.6840059>.
- [4] A Serener, S Serte. Geographic variation and ethnicity in diabetic retinopathy detection via deeplearning. *Turkish Journal of Electrical Engineering and Computer Sciences*, 2020, 28(2): 664-678.
- [5] B Zhang, Y Luo, L Ma, et al. 3D bioprinting: an emerging technology full of opportunities and challenges. *Bio-Design and Manufacturing*, 2018, 1(1): 2-13.
- [6] S Pathan, K G Prabhu, P Siddalingaswamy. Techniques and algorithms for computer aided diagnosis of pigmented skin lesions—A review. *Biomedical Signal Processing and Control*, 2018, 39: 237-262.
- [7] A Paradisi, S Tabolli, B Didona, et al. Markedly reduced incidence of melanoma and nonmelanoma skin cancer in a nonconcurrent cohort of 10,040 patients with vitiligo. *Journal of the American Academy of Dermatology*, 2014, 71(6): 1110-1116.
- [8] M E Celebi, Q Wen, H Iyatomi, et al. A state-of-the-art survey on lesion border detection in dermoscopy images. *Dermoscopy Image Analysis*, 2015: 97-129.
- [9] A Esteve, B Kuprel, R A Novoa, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 2017, 542(7639): 115.
- [10] A Steiner, H Pehamberger, K Wolff. Improvement of the diagnostic accuracy in pigmented skin lesions by epiluminescent light microscopy. *Anticancer Research*, 1987, 7(3): 433-434.
- [11] S Joseph, J R Panicker. Skin lesion analysis system for melanoma detection with an effective hair segmentation method. *2016 International Conference in Information Science (ICIS)*, IEEE, 2016: 91-96, <https://doi.org/10.1109/infosci.2016.7845307>.
- [12] P Zaenker, L Lo, R Pearce, et al. A diagnostic autoantibody signature for primary cutaneous melanoma. *Oncotarget*, 2018, 9(55): 30539.
- [13] C Barata, M Ruela, M Francisco, et al. Two systems for the detection of melanomas in dermoscopy images using texture and color features. *IEEE Systems Journal*, 2014, 8(3): 965-979.
- [14] T Vos, C Allen, M Arora, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: A systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*, 2016, 388(10053): 1545-1602.
- [15] P Wang, S Wang. Computer-aided CT image processing and modeling method for tibia microstructure. *Bio-Design and Manufacturing*, 2020, 3(1): 71-82.
- [16] Y LeCun, Y Bengio, G Hinton. Deep learning. *Nature*, 2015, 521(7553): 436.
- [17] Y LeCun, L Bottou, Y Bengio, et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [18] O Russakovsky, J Deng, H Su, et al. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 2015, 115(3): 211-252.
- [19] A Krizhevsky, I Sutskever, G E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 2012, 25: 1097-1105.
- [20] M D Zeiler, R Fergus. Visualizing and understanding convolutional networks. *European Conference on Computer Vision*, Springer, Cham, 2014: 818-833.
- [21] K Simonyan, A Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint [arXiv:1409.1556](https://arxiv.org/abs/1409.1556), 2014.
- [22] C Szegedy, W Liu, Y Jia, et al. Going deeper with convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015: 1-9. <https://doi.org/10.1109/CVPR.2015.7298594>.
- [23] K He, X Zhang, S Ren, et al. Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016: 770-778, <https://doi.org/10.1109/CVPR.2016.90>.
- [24] B Alipanahi, A DeLong, M T Weirauch, et al. Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning. *Nature Biotechnology*, 2015, 33(8): 831.
- [25] J Zhou, O G Troyanskaya. Predicting effects of noncoding variants with deep learning-based sequence model. *Nature Methods*, 2015, 12(10): 931.
- [26] A Shademan, R S Decker, J D Opfermann, et al. Supervised autonomous robotic soft tissue surgery. *Science Translational Medicine*, 2016, 8(337): 337ra64-337ra64.



- [27] S Kaymak, A Serener. Automated age-related macular degeneration and diabetic macular edema detection on OCT images using deep learning. *2018 IEEE 14th International Conference on Intelligent Computer Communication and Processing (ICCP)*, IEEE, 2018, <https://doi.org/10.1109/ICCP.2018.8516635>.
- [28] C Szegedy, V Vanhoucke, S Ioffe, et al. Rethinking the inception architecture for computer vision. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016: 2818-2826, <https://doi.org/10.1109/CVPR.2016.308>.
- [29] M Abadi, A Agarwal, P Barham, et al. TensorFlow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467, 2016.
- [30] Y Jia, E Shelhamer, J Donahue, et al. Caffe: Convolutional architecture for fast feature embedding. *Proceedings of the 22nd ACM International Conference on Multimedia*, 2014: 675-678, <https://doi.org/10.1145/2647868.2654889>.
- [31] F Bastien, P Lamblin, R Pascanu, et al. Theano: new features and speed improvements. arXiv preprint arXiv:1211.5590, 2012.
- [32] H Choi. Deep learning in nuclear medicine and molecular imaging: current perspectives and future directions. *Nuclear Medicine and Molecular Imaging*, 2018, 52(2): 109-118.
- [33] N Tajbakhsh, J Y Shin, S R Gurudu, et al. Convolutional neural networks for medical image analysis: Full training or fine tuning? *IEEE Transactions on Medical Imaging*, 2016, 35(5): 1299-1312.
- [34] Y Xu, T Mo, Q Feng, et al. Deep learning of feature representation with multiple instance learning for medical image analysis. *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2014: 1626-1630, <https://doi.org/10.1109/ICASSP.2014.6853873>.
- [35] E Long, H Lin, Z Liu, et al. An artificial intelligence platform for the multi-hospital collaborative management of congenital cataracts. *Nature Biomedical Engineering*, 2017, 1(2): 0024.
- [36] P Rajpurkar, J Irvin, K Zhu, et al. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint arXiv:1711.05225, 2017.
- [37] V Gulshan, L Peng, M Coram, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *Jama*, 2016, 316(22): 2402-2410.
- [38] S F Weng, J Rejs, J Kai, et al. Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS One*, 2017, 12(4): e0174944.
- [39] H C Hazlett, H Gu, B C Munsell, et al. Early brain development in infants at high risk for autism spectrum disorder. *Nature*, 2017, 542(7641): 348.
- [40] S Sarraf, G Tofghi. Classification of alzheimer's disease structural MRI data by deep learning convolutional neural networks. arXiv preprint arXiv:1607.06583, 2016.
- [41] N Amoroso, M La Rocca, S Bruno, et al. Brain structural connectivity atrophy in Alzheimer's disease. arXiv preprint arXiv:1709.02369, 2017.
- [42] L Rosado, M Ferreira. A prototype for a mobile-based system of skin lesion analysis using supervised classification. *2013 2nd Experiment International Conference (exp. at'13)*, IEEE, 2013: 156-157, <https://doi.org/10.1109/ExpAt.2013.6703051>.
- [43] J Hagerty, J Stanley, H Almubarak, et al. Deep learning and handcrafted method fusion: Higher diagnostic accuracy for melanoma dermoscopy images. *IEEE Journal of Biomedical and Health Informatics*, 2019: 1-1, <https://doi.org/10.1109/JBHI.2019.2891049>.
- [44] Andres, Diaz-Pinto, Sandra, et al. CNNs for automatic glaucoma assessment using fundus images: an extensive validation. *Biomedical Engineering Online*, 2019, 18(1), <https://doi.org/10.1186/s12938-019-0649-y>.
- [45] Y Li, L Shen. Skin lesion analysis towards melanoma detection using deep learning network. *Sensors*, 2018, 18(2): 556.
- [46] Y Gurovich, Y Hanani, O Bar, et al. Identifying facial phenotypes of genetic disorders using deep learning. *Nature Medicine*, 2019, 25(1): 60.
- [47] S S Han, M S Kim, W Lim, et al. Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm. *Journal of Investigative Dermatology*, 2018, 138(7): 1529-1538.
- [48] H Haenssle, C Fink, R Schneiderbauer, et al. Man against machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Annals of Oncology*, 2018, 29(8): 1836-1842, 2018.
- [49] C Mehanian, M Jaiswal, C Delahunt, et al. Computer-automated malaria diagnosis and quantitation using convolutional neural networks. *2017 IEEE International Conference on Computer Vision Workshop (ICCVW)*, IEEE, <https://doi.org/10.1109/ICCVW.2017.22>.
- [50] M Poostchi, K Silamut, R Maude, et al. Image analysis and machine learning for detecting malaria. *Translational Research the Journal of Laboratory & Clinical Medicine*, 2018, 194: 36-55.
- [51] Z I Attia, S Kapa, F Lopez-Jimenez, et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. *Nature Medicine*, 2019, 25(1): 70.
- [52] A Y Hannun, P Rajpurkar, M Haghpanahi, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature Medicine*, 2019, 25(1): 65.
- [53] J Zhang, Y Xie, Y Xia, et al. Attention residual learning for skin lesion classification. *IEEE Transactions on Medical Imaging*, 2019: 1-1, <https://doi.org/10.1109/TMI.2019.2893944>.
- [54] Y Fujisawa, Y Otomo, Y Ogata, et al. Deep-learning-based, computer-aided classifier developed with a small dataset of clinical images surpasses board-certified dermatologists in skin tumour diagnosis. *British Journal of Dermatology*, 2019, 180(61), <https://doi.org/10.1111/bjd.16924>.
- [55] A Rezvantab, H Safigholi, S Karimijeshni. Dermatologist level dermoscopy skin cancer classification using different deep learning convolutional neural networks algorithms. arXiv preprint arXiv:1810.10348, 2018.
- [56] K Yasaka, H Akai, A Kunimatsu, et al. Deep learning with convolutional neural network in radiology. *Japanese Journal of Radiology*, 2018: 1-16.
- [57] A Khamparia, P K Singh, P Rani, et al. An internet of health things-driven deep learning framework for detection and classification of skin cancer using transfer learning. *Transactions on Emerging Telecommunications Technologies*, 2020.
- [58] D Gutman, N C Codella, E Celebi, et al. Skin lesion analysis toward melanoma detection: A challenge at the international symposium on biomedical imaging (ISBI) 2016, hosted by the international skin imaging collaboration (ISIC). arXiv preprint arXiv:1605.01397, 2016.
- [59] L Bi, J Kim, E Ahn, et al. Automatic skin lesion analysis using large-scale dermoscopy images and deep residual networks. arXiv preprint arXiv:1703.04197, 2017.
- [60] S Serte, H Demirel. Gabor wavelet-based deep learning for skin lesion classification. *Computers in Biology and Medicine*, 2019, 113: 103423.
- [61] N C Codella, Q-B Nguyen, S Pankanti, et al. Deep learning ensembles for melanoma recognition in dermoscopy images. *IBM Journal of Research and Development*, 2017, 61(4/5): 5:1-5:15.
- [62] L Yu, H Chen, Q Dou, et al. Automated melanoma recognition in dermoscopy images via very deep residual networks. *IEEE Transactions on Medical Imaging*, 2017, 36(4): 994-1004.
- [63] X Fan, M Dai, C Liu, et al. Effect of image noise on the classification of skin lesions using deep convolutional neural networks. *Tsinghua Science and Technology*, 2020, 25(3): 425-434.
- [64] M Combalia, N Codella, V Rotemberg, et al. BCN20000: Dermoscopic Lesions in the Wild, arXiv preprint arXiv:1908.02288, 2019.
- [65] P Tschandl, C Rosendahl, H Kittler. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific Data*, 2018, 5(1): 1-9.
- [66] ISIC Project-ISIC Archive. Accessed: May 23, 2021. Available: <https://www.isic-archive.com>.
- [67] N Codella, D Gutman, M E Celebi, 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), 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), 2018: 168-172, <https://doi.org/10.1109/ISBI.2018.8363547>.
- [68] Y Yang, Y Ge, L Guo, et al. Development and validation of two artificial intelligence models for diagnosing benign, pigmented facial skin lesions. *Skin Research and Technology*, 2020, <https://doi.org/10.1111/srt.12911>.
- [69] Derm101 Image Library. Accessed: Jan. 12, 2019. Available: [https://www.derm101.com/image\\_library/](https://www.derm101.com/image_library/).
- [70] Dermnet-Skin Disease Atlas. Accessed: Dec. 31, 2018. Available: <http://www.dermnet.com/>.
- [71] H Mhaske, D Phalke. Melanoma skin cancer detection and classification based on supervised and unsupervised learning. *2013 International*

- Conference on Circuits, Controls and Communications (CCUBE), 2013: 1-5, <https://doi.org/10.1109/CCUBE.2013.6718539>.
- [72] I G Díaz. Incorporating the knowledge of dermatologists to convolutional neural networks for the diagnosis of skin lesions. *IEEE Journal of Biomedical and Health Informatics*, 2017, <https://doi.org/10.1109/JBHI.2018.2806962>.
- [73] O Abuzaghlh, B D Barkana, M Faezipour. Automated skin lesion analysis based on color and shape geometry feature set for melanoma early detection and prevention. *IEEE Long Island Systems, Applications and Technology (LISAT) Conference*, 2014: 1-6, <https://doi.org/10.1109/LISAT.2014.6845199>.
- [74] A Pennisi, D D Bloisi, D Nardi, et al. Skin lesion image segmentation using Delaunay Triangulation for melanoma detection. *Computerized Medical Imaging and Graphics*, 2016, 52: 89-103.
- [75] D D Gómez, C Butakoff, B K Ersboll, et al. Independent histogram pursuit for segmentation of skin lesions. *IEEE Transactions on Biomedical Engineering*, 2008, 55(1): 157-161.
- [76] S Kaymak, P Esmaili, A Serener. Deep learning for two-step classification of malignant pigmented skin lesions. *2018 14th Symposium on Neural Networks and Applications (NEUREL)*, 2018:1-6.
- [77] H Balazs. Skin lesion classification with ensembles of deep convolutional neural networks. *Journal of Biomedical Informatics*, 2018, 86: S1532046418301618-.
- [78] A Mahbod, G Schaefer, C Wang, et al. Transfer learning using a multi-scale and multi-network ensemble for skin lesion classification. *Computer Methods and Programs in Biomedicine*, 2020, 193: 105475.
- [79] A G Howard. Some improvements on deep convolutional neural network based image classification, arXiv preprint arXiv:1312.5402, 2013.
- [80] W Paja, M Wrzesień. Melanoma important features selection using random forest approach. *2013 6th International Conference on Human System Interactions (HSI)*, 2013: 415-418, <https://doi.org/10.1109/HSI.2013.6577857>.
- [81] F Nachbar, W Stolz, T Merkle, et al. The ABCD rule of dermatoscopy: High prospective value in the diagnosis of doubtful melanocytic skin lesions. *Journal of the American Academy of Dermatology*, 1994, 30(4): 551-559.
- [82] M Nasir, M Attique Khan, M Sharif, et al. An improved strategy for skin lesion detection and classification using uniform segmentation and feature selection based approach. *Microscopy Research and Technique*, 2018, 81(6): 528-543.
- [83] D G Lowe. Method and apparatus for identifying scale invariant features in an image and use of same for locating an object in an image: US, US6711293. 2004-3-23.
- [84] N Dalal, B Triggs. Histograms of oriented gradients for human detection. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005, 1: 886-893, <https://doi.org/10.1109/CVPR.2005.177>.
- [85] L Ballerini, R B Fisher, B Aldridge, et al. *A color and texture based hierarchical K-NN approach to the classification of non-melanoma skin lesions, color medical image analysis*. Dordrecht: Springer, 2013.
- [86] C Leo, V Bevilacqua, L Ballerini, et al. Hierarchical classification of ten skin lesion classes. *Proc. SICSA Dundee Medical Image Analysis Workshop*, 2015.
- [87] K Shimizu, H Iyatomi, M E Celebi, et al. Four-class classification of skin lesions with task decomposition strategy. *IEEE Transactions on Biomedical Engineering*, 2015, 62(1): 274-283.
- [88] A Zaidan, B Zaidan, O Albahri, et al. A review on smartphone skin cancer diagnosis apps in evaluation and benchmarking: Coherent taxonomy, open issues and recommendation pathway solution. *Health and Technology*, 2018: 1-16.
- [89] T-T Do, Y Zhou, H Zheng, et al. Early melanoma diagnosis with mobile imaging. *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, 2014: 6752-6757, <https://doi.org/10.1109/EMBC.2014.6945178>.
- [90] A Masood, A Al-Jumaily, K Anam. Self-supervised learning model for skin cancer diagnosis. *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, 2015: 1012-1015, <https://doi.org/10.1109/NER.2015.7146798>.
- [91] M F Duarte, T E Matthews, W S Warren, et al. Melanoma classification from Hidden Markov tree features. *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2012: 685-688, <https://doi.org/10.1109/ICASSP.2012.6287976>.
- [92] K Phillips, O Fosu, I Jouny. Mobile melanoma detection application for android smart phones. *2015 41st Annual Northeast Biomedical Engineering Conference (NEBEC)*, 2015: 1-2, <https://doi.org/10.1109/NEBEC.2015.7117184>.
- [93] F Topfer, S Dudorov, J Oberhammer. Millimeter-wave near-field probe designed for high-resolution skin cancer diagnosis. *IEEE Transactions on Microwave Theory & Techniques*, 2015, 63(6): 2050-2059.
- [94] I Valavanis, K Moutselos, I Maglogiannis, et al. Inference of a robust diagnostic signature in the case of Melanoma: Gene selection by information gain and Gene Ontology tree exploration. *13th IEEE International Conference on Bioinformatics and BioEngineering*, 2013: 1-4, <https://doi.org/10.1109/BIBE.2013.6701618>.
- [95] P Sabouri, H GholamHosseini, T Larsson, et al. A cascade classifier for diagnosis of melanoma in clinical images. *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014: 6748-6751, <https://doi.org/10.1109/EMBC.2014.6945177>.
- [96] M Efimenko, A Ignatev, K Koshechkin. Review of medical image recognition technologies to detect melanomas using neural networks. *BMC Bioinformatics*, 2020, 21(11): 1-7.
- [97] H L Semigran, D M Levine, S Nundy, et al. Comparison of physician and computer diagnostic accuracy. *Jama Intern. Med.*, 2016, 176(12): 1860-1861.
- [98] C Ross, I Swetlitz. IBM's Watson supercomputer recommended 'unsafe and incorrect' cancer treatments, internal documents show, *Stat News*, 2018, <https://www.statnews.com/2018/07/25/ibm-watson-recommended-unsafe-incorrect-treatments>.
- [99] D Castelvecchi. Can we open the black box of AI? *Nature News*, 2016, 538(7623): 20.
- [100] D Weinberger. Our machines now have knowledge we'll never understand, *Backchannel*, 2017, <https://www.wired.com/story/our-machines-now-have-knowledge-well-never-understand>.
- [101] A Körner, R Garland, Z Czajkowska, et al. Supportive care needs and distress in patients with non-melanoma skin cancer: Nothing to worry about? *European Journal of Oncology Nursing*, 2016, 20: 150-155.
- [102] O Malyuskin, V Fusco. Resonance microwave reflectometry for early stage skin cancer identification. *2015 9th European Conference on Antennas and Propagation (EuCAP)*, 2015: 1-6.
- [103] S Serte, A Serener, F Al-Turjman. Deep learning in medical imaging: A brief review. *Trans. Emerging Tel. Tech.*, 2020: e4080.
- [104] C M Doran, R Ling, J Byrnes, et al. Benefit cost analysis of three skin cancer public education mass-media campaigns implemented in New South Wales, Australia. *Plos One*, 2016, 11(1): e0147665.
- [105] A P Miller. Want less-biased decisions? Use algorithms. *Harvard Business Review*, 2018.
- [106] Gautam, Diwakar, Ahmed, et al. Machine learning-based diagnosis of melanoma using macro images. *International Journal for Numerical Methods in Biomedical Engineering*, 2018, 34(5): e2953.1.
- [107] W Fang, Y Li, H Zhang, et al. On the throughput-energy tradeoff for data transmission between cloud and mobile devices. *Information Sciences*, 2014, 283: 79-93, <https://doi.org/10.1016/j.ins.2014.06.022>.
- [108] J He, S L Baxter, J Xu, et al. The practical implementation of artificial intelligence technologies in medicine. *Nature Medicine*, 2019, 25(1): 30.
- [109] E J Topol. High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, 2019, 25(1): 44-56.