Ad Hoc Networks 80 (2018) 95–103

Contents lists available at ScienceDirect

Ad Hoc Networks

journal homepage: www.elsevier.com/locate/adhoc

DDA: A deep neural network-based cognitive system for IoT-aided dermatosis discrimination

Kehua Guo[a], Ting Li[a], Runhe Huang[b], Jian Kang[c,*], Tao Chi[d]

[a] School of Information Science and Engineering, Central South University, Changsha, China
[b] Faculty of Computer and Information Sciences, Hosei University, Tokyo, Japan
[c] Department of dermatology, the Third Xiangya Hospital, Central South University, Changsha, China
[d] Key Laboratory of Fisheries Information, Ministry of Agriculture, Shanghai Ocean University, Shanghai, China

ARTICLE INFO

Article history:
Received 31 January 2018
Revised 5 May 2018
Accepted 17 July 2018
Available online 18 July 2018

Keywords:
Deep neural network
Dermatosis discrimination
Cognitive system
Internet of things

ABSTRACT

The rapid development of the Internet of Things (IoT) and cognitive cyber-physical systems (CPS) has made people's daily lives more intelligent. Additionally, emerging technologies, such as wearable devices and machine learning, have demonstrated the potential for acquiring and processing large amounts of data from the physical world. In the medical field, effectively utilizing the collected medical data and providing more intelligent systems for doctors and patients to assist in diagnoses have also become important research topics. This paper presents a deep neural network-based cognitive system named DDA (dermatosis discrimination assistant) for classifying the dermatosis images generated by confocal laser scanning microscopes. Considering the lack of labels, we increase the labeled data automatically using an incremental model based on a small amount of labeled data and propose a disease discrimination model to distinguish and diagnose the categories of the disease images. In this system, the diagnoses of seborrheic keratosis (SK) and flat wart (FW) are used as examples, and experiments are conducted using the proposed models. Experimental results show that this system performs almost as well as individual dermatologists and can identify and diagnose other common dermatoses.

1. Introduction

The Internet of Things (IoT) and big data have been two most influential technologies of this era [1,2]. As breakthroughs in emerging technologies, the IoT, cloud computing and data analytics provide the potential to acquire and process large amounts of data from the physical world [3], thus making it possible to acquire, process and communicate data in intelligent applications [4]. Currently, these representative technologies have been applied in many fields, such as intelligent transportation systems, smart homes, smart buildings and smart cities [5,6]. Meanwhile, in the medical field, many scholars think that knowledge-based medical systems are a potential research topic.

Due to the development of IoT-based medical equipment, a large amount of medical data is constantly produced. Effectively analyzing and utilizing these data to provide more intelligent diagnostic tools to doctors and patients have become key issues. Recently, the deep convolutional neural network (CNN) has shown the potential to identify complex features of objects [7], and the prospect of handling medical big data has led to an interest in using deep learning to create professional medical robots with comprehensive abilities and high diagnostic accuracy [8,9].

In the medical field, the diagnoses of dermatoses or skin diseases mainly rely on doctors' visual observations and subjective experience and lack a scientific means of quantification [10]. For example, seborrheic keratosis (SK) and flat wart (FW) are two common skin diseases. Both often occur on the face, the back of the hand and the arm. These diseases present multiple lesions and affect the appearance [11]. Thus, these diseases may distress patients and seriously affect their physical and mental health. The clinical manifestation of SK is light brown macular or flat papules, and it is characterized as a smooth or slightly raised papilloma on the skin surface. Conversely, FW is characterized by a flat papule of unequal size with a slight uplift and smooth surface, which is rounded, elliptical or polygonal. It is difficult to distinguish the two diseases because their clinical manifestations are similar. Therefore, the accurate distinction between SK and FW is highly important for timely and effective treatment [12].

To better assist doctors in diagnosing the two common skin diseases, we propose a cognitive system named DDA (dermatosis discrimination assistant) for classifying medical images based on a deep convolutional neural network. The DDA can provide diagnostic advice to doctors. In general, the training in the deep con-

* Corresponding author.
E-mail address: kangjian_dm@126.com (J. Kang).

https://doi.org/10.1016/j.adhoc.2018.07.014
1570-8705/

volutional neural network requires large-scale data support. However, in the medical field, we inevitably encounter the problem of small data size and missing data labels. To address this problem, this paper uses a small amount of data labeled by doctors to automatically increase the labeled data and propose a discrimination model to distinguish and diagnose the category of the disease. Our dataset is constructed of images labeled by dermatology experts, unlabeled images and test images. The number of labeled images is small, and we use unlabeled images to increase the feature image dataset. We train the CNN using labeled images and the increased feature image dataset, which accurately identifies SK and FW.

The DDA is a cognitive system for IoT-aided dermatosis discrimination. On the one hand, in DDA, the CNN possesses the cognitive ability by learning the pathological knowledge from existing dermatological images and gives advice similar to that of a doctor using disease discrimination models. On the other hand, in practice, we can deploy the disease discrimination models on various IoT-based mobile terminals, such as mobile robots, lightweight equipments and sensors.

The remainder of this paper is organized as follows. Section 2 describes the related works. Section 3 describes the system model and key technologies, including the convolutional neural network, system work process, the dataset labeling incremental model and the performance evaluation method. Section 4 provides the experimental results. Section 5 presents the study's conclusions and outlines future work.

2. Related works

The skin is a natural barrier of the body and the first line of defense against external stimuli, such as ultraviolet light, detergents, mechanical friction and insect bites. Therefore, the incidence of skin diseases caused by environmental changes is rising. According to statistics, in the past several decades, cancer caused by skin diseases has appeared, and its serious consequences are self-evident. The traditional diagnosis of skin diseases is mostly based on clinical experience, which may lead to an incorrect diagnosis due to lack of experience or observation errors. Additionally, with the rapid increase in outpatient care, doctors need to observe a large number of cases and read massive amounts of image data, which increases the diagnostic workload. The occurrence of many diseases and the shortage of global medical resources create an urgent need to apply the knowledge-based auxiliary medical diagnosis system to hospitals [13].

Traditionally, research on the auxiliary diagnosis of dermatology has been conducted to automatically classify the images of pigmented skin lesions, and traditional image recognition methods are used to classify diseases [14]. Subsequently, to assist in the diagnosis of dermatologists, there have been many research results on computer-aided skin treatments for malignant melanomas [15], skin cancer [16] and other skin diseases [17]. Recently, new skin imaging technologies, such as skin testers, dermoscopes, and confocal laser scanning microscopes, have emerged, which advance the development of computer-aided systems for skin diseases. The massive number of skin images are collected from confocal laser scanning microscopes or skin professional medical equipment, such as dermoscopes. However, computer-aided diagnosis systems are usually not efficient because the skin disease images are complex due to the following factors: (1) the boundary of the lesion area is random, (2) the shape of the lesion is changeable, (3) the hair causes interference, and (4) the contrast of the images is dark and the feature extraction is difficult. Therefore, due to the similarities of some diseases' clinical features, the images are difficult to distinguish, and traditional image processing techniques are inefficient in classifying such images [18].

In the domain of medical imaging, the applications of machine learning, including computer-aided diagnosis, genomics and medical image analysis, are increasing. Recently, deep neural networks have made breakthroughs in the field of image processing, pattern recognition and computer vision [19]. Computer-aided diagnosis technology using deep learning methods has been used for several challenging tasks, such as health informatics [20] and medical image analysis [21]. Deep learning achieves many good results in the detection of mammographic lesions [22], the classification of pulmonary tuberculosis [23], and the segmentation of skin lesions [24]. In complex tasks, CNN has an unprecedented potential to improve the accuracy of classification and prediction [25]. For example, Reference [26] used deep learning to learn the characteristics of gene expression data to construct a classifier that detects and classifies different cancers. Reference [27] used a deep learning algorithm to assist doctors in diagnosing potential patients with congenital cataracts using identification networks, evaluation networks and strategist networks.

However, the potential of deep neural networks has not been fully demonstrated to date in the field of medical imaging. To achieve better results in disease detection and classification, deep learning requires a labeled dataset. Dataset labeling often requires professionals, and this labeling is a highly time-consuming and expensive job. Increasing the labeled data when there is a small amount of labeled data to allow for the effective use of deep learning methods has become an issue that researchers are constantly exploring. Thus, several semi-supervised learning methods for a small portion of labeled data have been proposed, such as greedy algorithms [28] and embedded learning [29], which require a large number of computations. In DDA, we expand the training dataset based on a small part of labeled data using an incremental labeling model**.** The DDA completes the tests that were comparable to the dermatologists and demonstrates the ability to identify and classify the SK and FW at a level comparable to the doctors. The proposed system can be extended to identify and diagnose other common dermatoses.

3. Methodologies in DDA

3.1. Convolutional neural network

The DDA uses CNN for dermatosis discrimination. We exploit the GoogleNet Inception v3 [7] model architecture, which is pre-trained using 1.28 million images (1000 object categories) from the ImageNet Large Scale Visual Recognition Challenge of 2014 [30], and we train it on our dataset. Fig. 1 shows the dermatosis discrimination using GoogleNet Inception v3.

The GoogleNet Inception v3 model has 46 layers. The model includes 11 Inception modules. In Fig. 1, the inception constructed in the GoogleNet Inception v3 model combines different convolution layers in parallel. In this paper, we directly train the disease discrimination model with images generated by confocal laser scanning microscope images and use the pixel data and disease labels as input.

The disease dataset includes labeled data and unlabeled data, and the labeled data include both feature data and non-feature data. The feature data contain images with obvious and clear lesion characteristics. Correspondingly, the non-feature data contain images with no obvious lesion features and noise information. Additionally, our dataset also contains some image data that are not labeled, and we need to label them with a semi-supervised learning algorithm. We assume a dermatosis dataset containing the images from N skin diseases that are represented as $DS=\{DS_1,DS_2,\ldots, DS_N\}$ and give the definition of the dataset as follows.

Definition 2.1. $\forall DS_i \in DS$ and $DS_i=\{A_{il},A_{iu}\}$ satisfy

$$\begin{cases} A_{il}=\{S_{il},T_{il}\} \\ A_{iu}=\left\{a_{i1},a_{i2},\ldots,a_{|A_{iu}|}\right\} \end{cases} \quad (1)$$

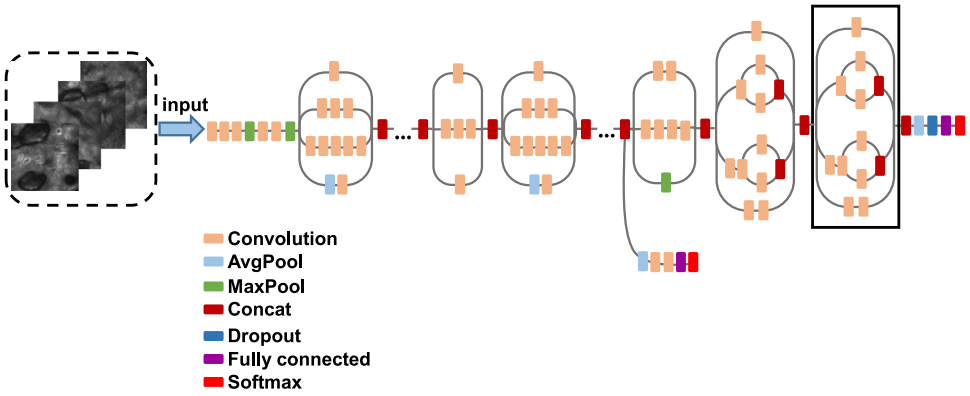

Fig. 1. Dermatosis discrimination using GoogleNet Inception v3.

where

(1) A_{il} and A_{iu} are the labeled and unlabeled image set, respectively,
(2) S_{il} is a labeled feature image set and $S_{il} = \{s_{i1}, s_{i2}, \ldots, s_{|S_{il}|}\}$, and
(3) T_{il} is a labeled non-feature image set and $T_{il} = \{t_{i1}, t_{i2}, \ldots, t_{|T_{il}|}\}$.

There are three different layers in CNN, including the convolutional layers, pooling layers and fully connected layers. In this paper, we pre-train Google's Inception v3 CNN architecture using the ImageNet training dataset. Next, the last classification layer of the network is removed and retrained using our dataset. For the discrimination of N skin diseases, when an image I is input into the trained model, the target feature vector $Z=(z_1,z_2,\ldots, z_N)$ is output after the feature extraction process and the last fully connected layer training. Next, the probability is calculated using the function of the softmax layer to classify the labeled images. We use the same method to classify the unlabeled images, and the probability calculation formula and classification formula are as follows.

Definition 2.2. $\forall I$ is an unlabeled image. The classification of I can be expressed as follows.

$$\begin{cases} h(I) = softmax(Z) = [p(q=1|\ I), p(q=2|\ I), \ldots, p(q=N|\ I)]^T \\ class(I) = i,\ \ when\ \ p(q=i|\ I) = argmax(p(q=j|\ I))\, for\ \ j = 0 \ldots N \end{cases} \tag{2}$$

3.2. System work process

The system identifies the categories of N skin diseases from a large number of patients based on a small amount of labeled data. The categories of the diseases are identified using the disease discrimination model, which provides a reference for the treatment of skin diseases in patients and helps dermatologists in decision-making. The dermatologists provide the final treatment decision according to the results provided by the disease discrimination model and their observations. In DDA, the diagnoses of seborrheic keratosis (SK) and flat wart (FW) are used as examples. The identification process is shown in Fig. 2.

As shown in Fig. 2, the work process of the system consists of five steps. (1) The dermatology experts label some images to obtain A_{il} and A_{iu}. (2) A_{il} is trained independently to obtain the incrementally labeled data model. (3) A_{iu} is identified and labeled using the incremental model, and the resulting A_{il} can make the training dataset extension. (4) The disease discrimination model can be generated by training the labeled A_{il} using the incremental model. In practice, the disease discrimination model can be distributedly deployed on various mobile terminals. (5) Users input the test image, and the disease discrimination model can identify it and give the test result.

3.3. Incremental dataset labeling model

Our training dataset is obtained from the Xiangya dermatology database, which includes the images from N skin diseases. Experienced dermatology experts have described and labeled a part of the images independently and generated S_{il} and T_{il}. The remaining image set in A_{iu} is unlabeled. It contains both feature images and non-feature images, and the non-feature images will affect the accuracy rate. We use an incremental model to automatically classify the unlabeled data to extend the scale of A_{il}. Therefore, $\forall DS_i \in DS$, DS_i consists of the following parts: (1) $|S_{il}|$ feature images and $|T_{il}|$ non-feature images, which are labeled by dermatology experts and trained using deep convolutional neural networks; and (2) $|A_{iu}|$ unlabeled images.

The dataset used in the DDA includes only a small amount of labeled data, while the rest of the data are unlabeled. In the incremental model, our aim is to increase the number of labeled images in A_{il}. In DS_i, we divide the unlabeled dataset A_{iu} into α parts, which can be represented as $A_{iu} = \{A_{iu1}, A_{iu2}, \ldots, A_{iu\alpha}\}$. Then, one part in A_{iu} is selected to be input into the incremental model for the corresponding disease to be labeled. The incremental dataset labeling algorithm is represented as Algorithm 1.

In Algorithm 1, we use feature images and non-feature images labeled by dermatology experts to train the first incremental model. After these images are labeled as feature or non-feature, they will be trained again with the previous training dataset to update the incremental model. Thus, one part of A_{iu} will be selected again from the unlabeled dataset as an input to the second incremental model for labeling. Based on this method, the experiment is conducted α times, and the feature images are increased to a certain number.

Algorithm 1 has to conduct α trainings, which is computationally expensive. To solve this problem, this process can be performed in a background thread. Generally, the data testing takes less time than training. Therefore, calculating class (I) using model (k) and labeling class (I) to I can be finished in a short time. Hence, besides the training time, the computation complexity of the incremental dataset labeling algorithm is O($|DS|$) (where $|\#|$ represents the cardinality of a set).

3.4. Performance evaluation model

3.4.1. Naive identification and K-fold cross-validation method

This method is proposed based on the incremental dataset labeling model. To validate the effectiveness, we train the disease discrimination model using the labeled data and the incremental

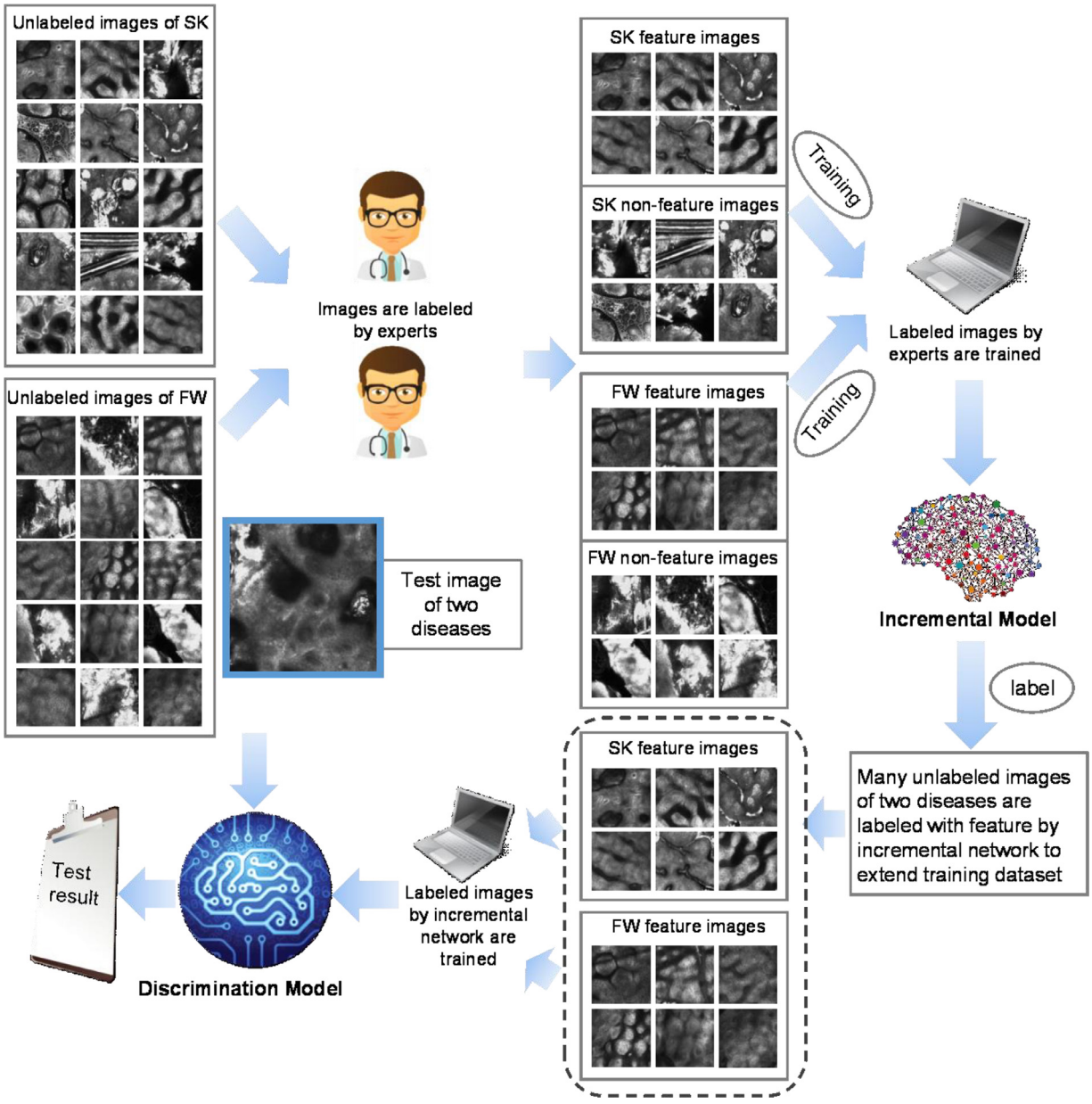

Fig. 2. System work process.

Algorithm 1: Incremental dataset labeling algorithm.

```
Input: DS={DS1,DS2,..., DSN};
for k=1 to α do
        A=∪_{i=1}^{N} A_il;
        Train (A)→generate incremental model(k);
        for each image I in A_iuk do
          calculate class(I) using model(k);
        label class(I) to I;
        if I is a feature image then
          add I to S_ilk;
        else
          add I to T_ilk;
        end if
      end for
end for
```

data and use the test dataset $DT=\{c_1,c_2,\ldots, c_h\}$ to validate the result.

In naive identification, we use $S_{il}'=\{S_{il1},S_{il2},\ldots, S_{il\alpha}\}$ to define the feature image set labeled by each incremental model and add the incremental feature image set S_{il}' to the labeled feature image datasetS_{il}. Therefore, S_{il} includes $\alpha+1$ parts, which can be represented as $S_{il}=\{S_{il0},S_{il1},\ldots, S_{il\alpha}\}$, in which S_{il0} is the feature image dataset of all diseases labeled by dermatology experts. The first training dataset is S_{il0}, and we use the generated model to test DT and obtain an accuracy rate r_0. Next, we train $S_{il0}\cup S_{il1}$, generate a new model to test DT, and obtain an accuracy rate r_1. This step is repeated until the dataset labeling using the incremental model finishes. We obtain a vector $r=\{r_0,r_1,r_2...,r_\alpha\}$. The naive identification algorithm is presented as Algorithm 2.

The complexity of Algorithm 2 is similar to Algorithm 1. In addition to the training time, calculating class (I) using model (k) can be finished in $O(|DT|)$ time.

The K-fold cross method is applied to verify the independence. All the data labeled using the incremental model is randomly divided into K parts. Among the K parts, one is selected as the test data, and the remaining are used as the training data. The step repeats K times, and in each iteration, a different part is selected from the K parts as the test data to ensure that all K parts of the

Algorithm 2: Naive identification algorithm.

```
Input: DS = {DS_1, DS_2, ..., DS_N}, DT = {c_1, c_2, ..., c_h};
Output: r = {r_0, r_1, r_2 ..., r_α};
S = ∅;
for k = 0 to α do
    S = S ∪ (∪_{i=1}^{N} S_ilk);
  Train (S) → generate incremental model(k);
    for each image I in DT do
      calculate class(I) using model(k);
    end for
     calculate r_i;
end for
```

data are test data. Finally, we can obtain K results and get the average value. We should ensure that the training data are not used for testing, and the training model is preserved before any validation.

3.4.2. Multi-clinical patient trial

The third hospital of Xiangya participates in the practical verification of the disease discrimination model, and we select some images of eligible patients for testing. The patient's image is uploaded to the disease discrimination model, and we obtain the result from the disease discrimination model.

The dermatology experts design some cases of confocal laser scanning microscope images involving various clinical settings. This method compares the performance of the disease discrimination model with the dermatologist's performance. Each case is identified as a skin disease. In this method, the diagnoses of SK and FW are used as examples, and three dermatologists with expertise (experts, dermatology attending doctor and novice) are asked to independently assess the same test images with the disease discrimination model without any other information. The expert is a professor with many years of experience in a dermatology department. The dermatology attending doctor is a doctor who completed special training in clinical training and dermatology. The novice doctor is a student who completed dermatological research and started clinical practice.

4. Experimental results

4.1. Incremental dataset labeling experiment

In this paper, the diagnoses of SK and FW are taken as examples. We collect 755 SK images and 545 FW images, which include 43 labeled feature images and 48 labeled non-feature images of SK, and 44 labeled feature images and 41 labeled non-feature images of FW. The incremental experiment is a process of automatic labeling by machine. In this process, our goal is to label the feature images to expand the training data set. To ensure the experimental performance, the number of images cannot be too small in each increment of training. Meanwhile, considering the number of labeled feature images, we cannot add too many images to the incremental training. The unlabeled images include 664 images of SK and 460 images of FW. We divide the two datasets into four parts, and each time we select one part as the testing dataset to label. To simplify the experimental environment, the machine used for training and testing is MacBook Pro (2.7 GHz Intel Core i5 processor, 8GB memory). In fact, we can deploy the disease discrimination models on various IoT based mobile terminals such as mobile robots, lightweight equipments and sensors. We record the results of the 166 unlabeled images of SK in each test, show the results of our four incremental dataset labeling experiments and record the results of the 115 unlabeled images of FW in each test. The incremental results are shown in Table 1.

As shown in Table 1, the increment no. column indicates the number of training iterations, and the training data column represents the set of training data. Each training dataset includes feature images and non-feature images. The training dataset in the first iteration contains images labeled by dermatology experts, and the training dataset of each iteration is composed of the previous training dataset and the previous result set. The test data column in Table 1 shows the test data, and the last two columns in the table show the test results, which are the incremental feature images and non-feature images. Table 1 shows that the number of feature images increases consistently in the first four iterations. The feature images of the two diseases are increased to 356 $(43+62+87+83+81)$ for SK and 399 $(44+105+92+91+67)$ for FW. The experiment is performed based on these labeled data.

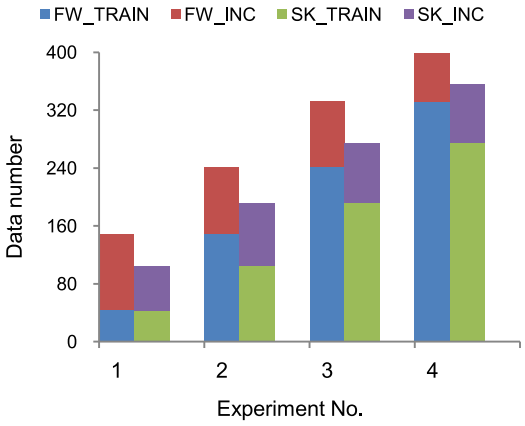

Fig. 3. Number of training data.

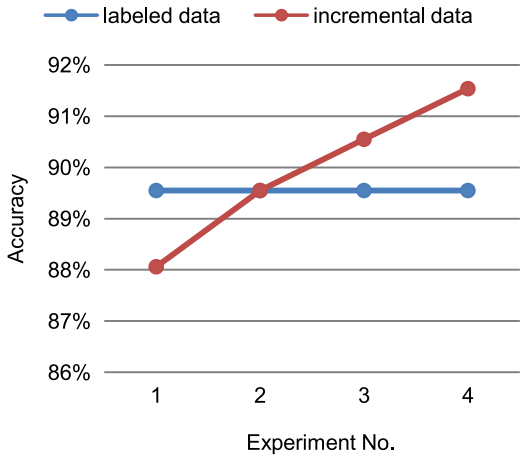

Fig. 4. Results of naive identification.

4.2. Naive identification experiment

This experiment is based on labeling the dataset incrementally. We conduct four training sessions, and each training data are feature images of the incrementally labeled data. In this experiment, 201 mixed images of SK and FW are selected as the test dataset. The four disease discrimination models are tested with the 201 mixed images. Fig. 3 shows the number of data used for each training and testing period, which is composed of training data and the result sets in the incremental dataset labeling. The accuracy of each model is recorded in Fig. 4. The accuracy is obtained by testing each model with the test data.

The accuracy of the model with the images labeled by dermatology experts is 89.55%, and the accuracies of the model with incremental data are 88.06%, 89.55%, 90.55% and 91.54%, respectively. Fig. 4 shows that the accuracy of the first disease discrimination model with incremental data is lower than that of the model with

Table 1
Results of the SK and FW incremental dataset labeling experiment.

	Increment no.	Taining data(feature images/non-feature images)	Test data	Incremental feature image	Incremental non-feature image
Seborrheic Keratosis (SK)	1	43/48	166	62	104
	2	105/152	166	87	79
	3	192/231	166	83	83
	4	275/314	166	81	85
Flat Wart (FW)	1	44/41	115	105	10
	2	149/51	115	92	23
	3	241/74	115	91	24
	4	332/98	115	67	48

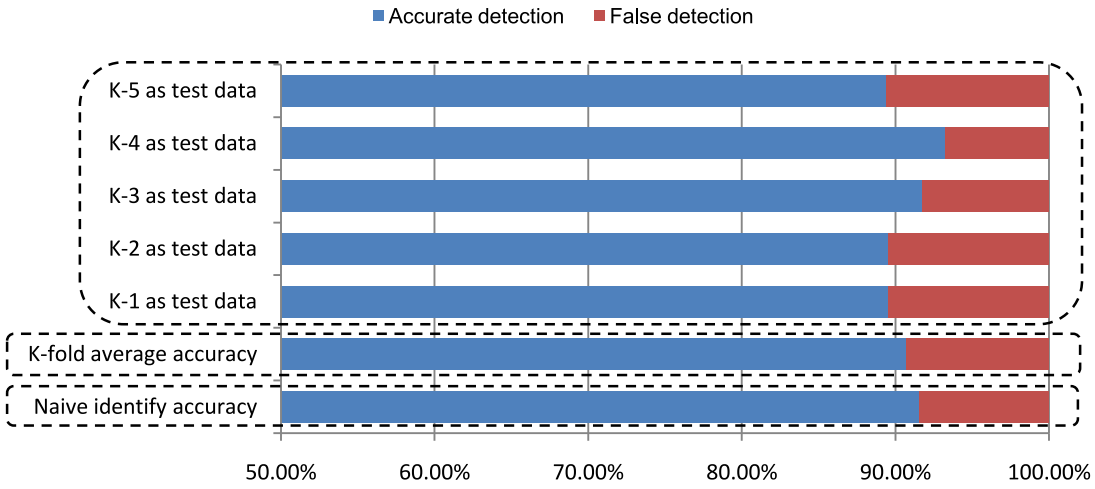

Fig. 5. Results of *K*-fold cross-validation test.

the labeled images. This is because the labeled images are all feature images labeled by dermatology experts, and the first training data labeled by the incremental model may contain some images without obvious features, which reduces the recognition accuracy. However, as the training data increase, the identification accuracy is also rising.

These results indicate that the incremental data are correct and effective and also demonstrate that we could increase the amount of data with the deep learning algorithm when there is limited labeled data. Meanwhile, it can also be seen from the results that the greater the amount of training data, the higher the accuracy of the disease discrimination model.

In the training processes, we use ordinary servers, and each round of training takes approximately 5 minutes, which is a relatively long time. However, if we switch the training to high-performance servers, the training time will be greatly reduced. We use an ordinary terminal to test the disease identification model, and it only takes 1–2 s to complete the identification of SK and FW. If we deploy the model to IoT terminals, this time is acceptable. In our paper, the innovation is an algorithmic framework, and the focus is to make the machine possess the cognitive ability to assist doctors in the diagnosis rather than to exceed the speed of doctors' discrimination.

4.3. K-fold cross-validation

Cross-validation experiments are conducted to verify the independence and validity of the data obtained by incremental labeling. On the one hand, considering that the total number of feature images obtained by the incremental experiment and the small number of images in each test may affect the accuracy, the K value should not be too large. On the other hand, if the K value is too small, the number of tests will be too small, and the average value is not accurate enough. Therefore, we use K-fold cross-validation (K = 5) to test the data obtained in the incremental dataset labeling. We divide the data obtained from the incrementally labeled dataset into five parts, which are denoted by E_1, E_2, E_3, E_4 and E_5. Each one is taken as testing data, and the remaining parts are used as the training data. We record the results of each experiment with Acc_{E_i} and record the average accuracy with *Avg_acc*.The formula for calculating the average accuracy is as follows.

$$Avg\_acc = \frac{1}{K}\sum_{i=1}^{K} Acc_{E_i} \tag{3}$$

The accuracy of each experiment and the average accuracy are shown in Fig. 5. In Fig. 5, the naive identification accuracy is the last result in the naive identification experiment.

Fig. 5 shows that the accuracy of each experiment is high, and the average accuracy reaches 90.72%. The results of this experiment show that the incremental data in the incrementally labeled dataset are correct and independent. Regardless of which data are used as the training dataset, the accuracy of the recognition is very high and comparable to the naive identify accuracy.

4.4. Multi-clinical patient diagnosis

The data used in the above experiments are collected from the Xiangya dermatology database. To further investigate the versatility and utility of the disease discrimination model, we conduct a multi-clinical patient trial and test using non-specialist hospital cases. We select 52 patients from January 2015 to March 2017 from the third Hospital of Xiangya and select an image from each patient's medical record as the testing data. The expert and the disease discrimination model identify the 52 images independently. The result of each recognition is recorded, as shown in Fig. 6.

Fig. 6 shows the diagnosis results by a dermatology expert and the disease discrimination model, and the red triangle marks the

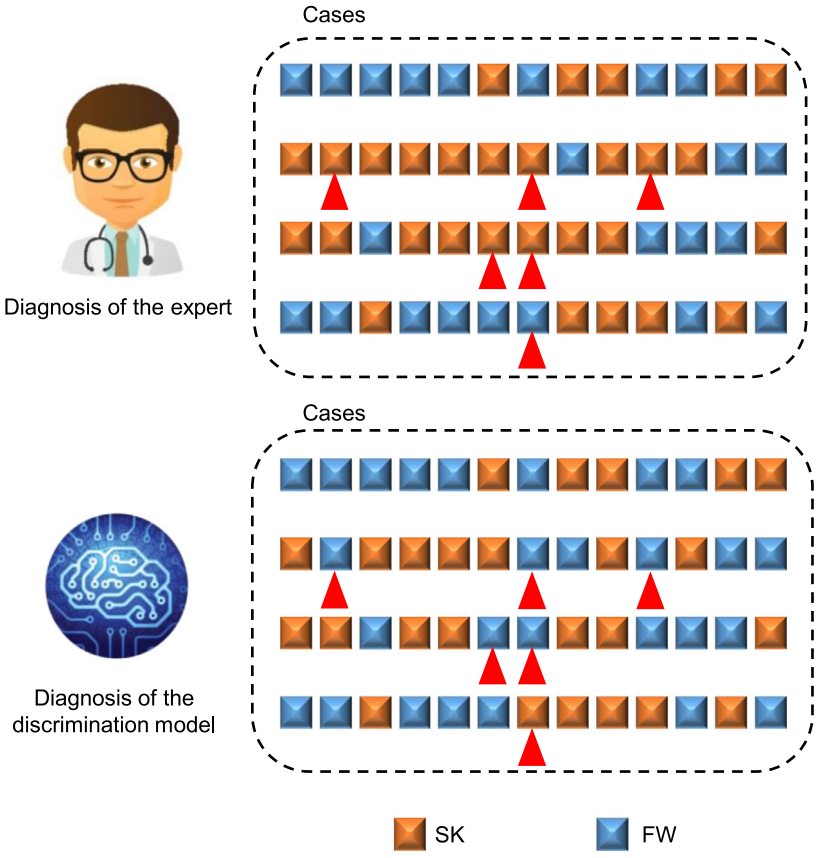

Fig. 6. Results of the multi-clinical patient test.

different diagnostic results of the expert and the disease discrimination model. With the expert's diagnosis as a reference, we can see that the disease discrimination model identifies 6 errors in 52 images. Therefore, the accuracy of the disease discrimination model reaches 88.5% in the clinical experiment, which indicates its versatility and utility.

4.5. Contrast test

We compare the results of the expert group and the disease discrimination model. To be more realistic, we conduct a test to compare the performance of the disease discrimination model with the medical level of the expert group. For the purpose of assessing the real-world performance of the disease discrimination model in clinical practice, a panel of experts designed a test document, which consists of 109 confocal laser scanning microscope images that meet the clinical requirements. Three dermatologists with different degrees of expertise (expert, dermatology attending doctor and novice) and the disease discrimination model complete the same test independently. The test results are shown in Fig. 7.

In Fig. 7, the red box indicates the incorrect identification. The results in Fig. 7 show that the expert identifies 3 out of 109 cases incorrectly. The dermatology attending doctor identifies 5 cases incorrectly. The novice doctor identifies 12 cases incorrectly. The disease discrimination model identifies 7 out of 109 cases incorrectly. The disease discrimination model performs well compared to the dermatologist, which is comparable to the performance of the dermatology attending doctors. In this test, the disease discrimination model provides an accurate identification of these cases. We believe that the performance of the system is comparable to that of a qualified dermatologist.

5. Discussion and conclusion

Previously, skin disease research focused on image processing technology to identify the categories of diseases. There are limited studies on the identification and classification of SK and FW. In this paper, confocal laser scanning microscope images with varying degrees of SK and FW are used. However, the machine and the doctor are different, and the machine creates differences in the image's light intensity, shooting depth and angle. This interference makes the image more complex in our experiments and increases the challenge of identification. However, in all of the tests, the high accuracy of the disease discrimination model shows that the DDA can robustly distinguish the categories of SK and FW in the real world. This highlights the powerful potential and the advantages of deep learning algorithms in identifying complex images in realistic clinical challenges.

The rapid development of artificial intelligence and deep learning has been achieved using big data [30]. For SK and FW, it is difficult to obtain a large amount of high-quality data for research, although these two skin diseases are common. This difficulty is observed because the isolation of the medical data is a bottleneck in data usage, and accurate labeling of medical data is also cumbersome and requires professional experience. However, the method of incrementally labeling datasets proposed in this paper can solve this problem. It uses limited labeled images to train an incremental model. Next, the incremental model identifies and labels the unlabeled images to increase the amount of training data. The result of the naive identification test and K-fold cross-validation show that the accuracy of the disease discrimination model is improved after increasing the training data. This finding demonstrates that increasing the amount of data using the incremental model is useful. If more data are collected and integrated, the accuracy of the disease discrimination model will increase.

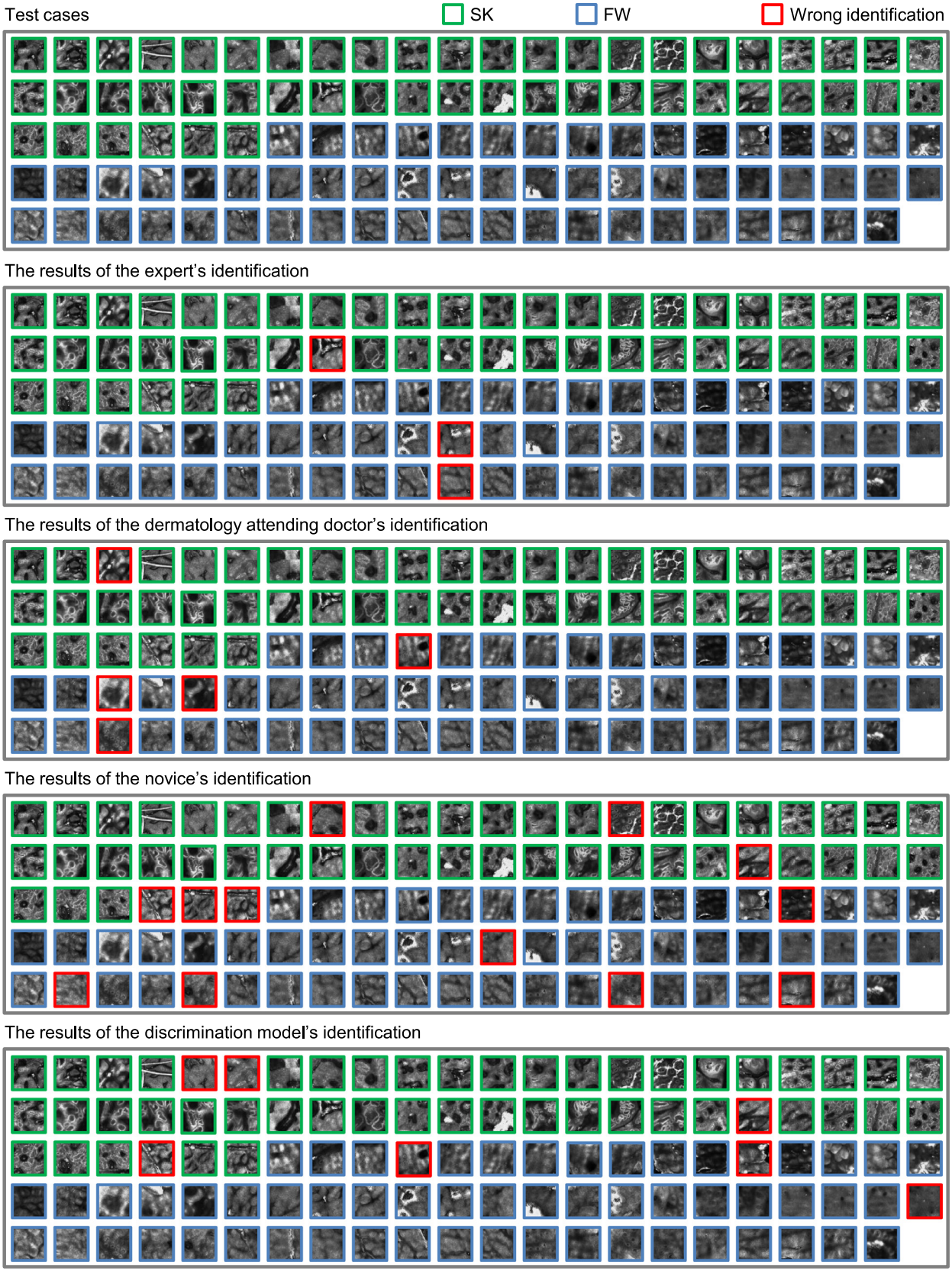

Fig. 7. Results of the contrast test.

The high accuracy of the DDA is realized in specific situations. Our next step is to validate whether the DDA can be extended to the identification and classification of other skin diseases in different clinical circumstances. Simultaneously, we will explore the feasibility of the system in clinical practice by allowing the cognitive system to be used for clinical auxiliary diagnosis.

Acknowledgments

This work is supported by Natural Science Foundation of China (61672535, 61502540), Key Laboratory of Information Processing and Intelligent Control of Fujian Innovation Fund (MJUKF201735). The authors declare that they have no conflict of interests.

References

[1] X. Zhou, W. Wang, Q. Jin, Multi-dimensional attributes and measures for dynamical user profiling in social networking environments, Multimed. Tools Appl. 74 (14) (2015) 5015–5028.
[2] Y. Zhang, J. Ren, J. Liu, et al., A survey on emerging computing paradigms for big data, Chin. J. Electron. 26 (1) (2017) 1–12.
[3] X. Wang, L.T. Yang, J. Feng, et al., A tensor-based big service framework for enhanced living environments, IEEE Cloud Comput. 3 (6) (2016) 36–43.
[4] N. Zhong, J. Ma, R. Huang, et al., in: Research Challenges and Perspectives on Wisdom Web of Things (W2T), Wisdom Web of Things. Springer International Publishing, 2016, pp. 3–26.
[5] X. Wang, L.T. Yang, X. Xie, et al., A cloud-edge computing framework for cyber-physical-social services, IEEE Commun. Mag. 55 (11) (2017) 80–85.
[6] X. Zhou, Q. Jin, B. Wu, et al., Organic streams: data aggregation and integration based on individual needs, in: 2013 International Joint Conference on Awareness Science and Technology and Ubi-Media Computing, IEEE, 2013, pp. 535–541.
[7] A. Esteva, B. Kuprel, A. Novoa R, et al., Dermatologist-level classification of skin cancer with deep neural networks, Nature 542 (7639) (2017) 115–118.
[8] Y. Bar, I. Diamant, L. Wolf, et al., Chest pathology identification using deep feature selection with non-medical training, Comput. Methods Biomech. Biomed. Eng. (2016) 1–5.
[9] M. Turan, Y. Almalioglu, E. Konukoglu, et al., A Deep Learning Based 6 Degree-of-Freedom Localization Method For Endoscopic Capsule Robots. arXiv preprint arXiv:1705.05435, 2017.
[10] N.J. Dhinagar, M. Celenk, Early diagnosis and predictive monitoring of skin diseases, in: Healthcare Innovation Point-Of-Care Technologies Conference, IEEE, 2016, pp. 33–36.
[11] K.P. Kyriakis, G. Pagana, C. Michailides, et al., Lifetime prevalence fluctuations of common and plane viral warts, J. Eur. Acad. Dermatol. Venereol. 21 (2) (2007) 260.
[12] J.M. Yeatman, M. Kilkenny, R. Marks, The prevalence of seborrhoeic keratoses in an Australian population: does exposure to sunlight play a part in their frequency? Br. J. Dermatol. 137 (3) (1997) 411–414.
[13] B. Farahani, F. Firouzi, V. Chang, et al., Towards fog-driven IoT eHealth: promises and challenges of IoT in medicine and healthcare, Future Gener. Comput. Syst. 78 (2018) 659–676.
[14] I. Zaqout, Diagnosis of skin lesions based on dermoscopic images using image processing techniques, Int. J. Signal Process. Image Process. Pattern Recognit. 9 (9) (2016) 189–204.
[15] A. Mikołajczyk, A. Kwasigroch, M. Grochowski, Intelligent system supporting diagnosis of malignant melanoma, in: 2017 Polish Control Conference, Springer, Cham, 2017, pp. 828–837.
[16] H. Castillejos-Fernándeza, O. López-Ortegaa, F. Castro-Espinozaa, et al., An intelligent system for the diagnosis of skin cancer on digital images taken with dermoscopy, Acta Polytech. Hung. 14 (3) (2017) 169–185.
[17] D.M. Khairina, H.R. Hatta, R. Rustam, et al., Automation diagnosis of skin disease in humans using dempster-shafer method, in: E3S Web of Conferences, 31, EDP Sciences, 2018, p. 11006.
[18] A.A. Zaidan, B.B. Zaidan, O.S. Albahri, et al., A review on smartphone skin cancer diagnosis apps in evaluation and benchmarking: coherent taxonomy, open issues and recommendation pathway solution, Health Technol. (2018) 1–16.
[19] H.C. Shin, H.R. Roth, M. Gao, et al., Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning, IEEE Trans. Med. Imag. 35 (5) (2016) 1285–1298.
[20] D. Ravì, C. Wong, F. Deligianni, et al., Deep learning for health informatics, IEEE J. Biomed. Health Inform. 21 (1) (2017) 4–21.
[21] D. Shen, G. Wu, I. Suk H, Deep learning in medical image analysis, Ann. Rev. Biomed. Eng. 19 (2017) 221–248.
[22] T. Kooi, G. Litjens, B. van Ginneken, et al., Large scale deep learning for computer aided detection of mammographic lesions, Med. Image Anal. 35 (2017) 303–312.
[23] P. Lakhani, B. Sundaram, Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks, Radiology 284 (2) (2017) 574–582.
[24] Y. Yuan, M. Chao, Y.C. Lo, Automatic skin lesion segmentation using deep fully convolutional networks with Jaccard distance, IEEE Trans. Med. Imag. 36 (9) (2017) 1876–1886.
[25] K. He, X. Zhang, S. Ren, et al., Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2016, pp. 770–778.
[26] R. Fakoor, F. Ladhak, A. Nazi, et al., Using deep learning to enhance cancer diagnosis and classification, in: Proceedings of the International Conference on Machine Learning, 2013.
[27] E. Long, H. Lin, Z. Liu, et al., An artificial intelligence platform for the multihospital collaborative management of congenital cataracts, Nat. Biomed. Eng. 1 (2017) 0024.
[28] D. Xiao, R. Wang, L. Wu, Latent attribute inference of users in social media with very small labeled dataset, IEICE Trans. Inf. Syst. 99 (2016) 2612–2618.
[29] E. Hoffer, N. Ailon, Semi-supervised Deep Learning By by Metric Embedding. arXiv preprint arXiv:1611.01449, 2016.
[30] O. Russakovsky, J. Deng, H. Su, et al., Imagenet large scale visual recognition challenge, Int. J. Comput. Vis. 115 (3) (2015) 211–252.

Kehua Guo is a professor at the School of Information Science and Engineering, Central South University, China. He received the Ph.D. degree in Computer Science and Technology from Nanjing University of Science and Technology. His research interests include artificial intelligence and, ubiquitous computing.

Ting Li received the B.SC. degree in Computer Science and Technology from Xiangtan University in 2016. She is currently a graduate student at the School of Information Science and Engineering, Central South University, China. Her research interests include ubiquitous computing and big data.

Runhe Huang received Ph.D. in Computer Science and Mathematics from the University of the West of England. She is a professor in the Faculty of Computer and Information Sciences at Hosei University, Japan. Her research fields include computational intelligence, and ubiquitous intelligence computing.

Jian Kang received the Ph.D. degree in Medicine from the Nanjing Medical University, China. He is a doctor in the Department of Dermatology of the Third Xiangya Hospital of Central South University, China. His research fields include information and network management and diagnosis of dermatosis.

Tao Chi received the Ph.D. degree in Tongji University, China. He is a professor in the Key Laboratory of Fisheries Information, Ministry of Agriculture, Shanghai Ocean University. His research fields include Internet of Things and ubiquitous computing.