Original Article
Study on deep convolutional neural network-learning vector quantization model and its application in breast adenosis prediction based on infrared thermography

Zhen Long[1], Zichen Lu[2], Longlong Liu[1], Xiaoyan Lu[3]

[1]The School of Mathematical Sciences, Ocean University of China, Qingdao 266000, PR China; [2]The Cloud Computing Center, Chinese Academy of Science, Dongguan 523000, PR China; [3]The Dongguan Hospital of Traditional Chinese Medicine, Dongguan 523000, PR China

Abstract: This study proposes a reliable and efficient method for analyzing medical infrared thermography (IRT) images in order to provide a more accurate and easier diagnosis of breast adenosis in patients. The proposed method is based on an improved deep convolutional neural network model named the DCNN-LVQ, which represents a combination of the deep convolutional neural network (DCNN) and learning vector quantization (LVQ) algorithm. Image features are first extracted from the deep convolutional neural network, and then classified by the learning vector quantization algorithm. The proposed method was verified with the experiment using the infrared thermography images of patients with breast adenosis and healthy persons, which were labeled by medical experts. The disease prediction accuracies reached 97.52% and 94.66% respectively for the two sample datasets, which could be considered a satisfactory result. The proposed model was verified experimentally and was compared with three common prediction models. The results showed that the proposed model was superior to the other models in infrared thermography image recognition. Thus, the proposed model provide breast adenosis prediction with higher accuracy. In addition, the proposed model is highly robust, making it high likely to be applied to other infrared thermography image recognition problems.

Keywords: Convolutional neural network, infrared thermography, learning vector quantization

Introduction

According to the report of the China Population Association in Beijing in February 2010, the mortality rate from breast cancer in Chinese cities has increased by 38.91% from 2003 to 2009. The incidence of breast cancer ranks first among the tumors in women in big cities and it has become the most threating disease to women's health for decades [1]. In the asymptomatic female population, the incidence of various breast diseases has reached 52.4%, which is a much higher than those of other chronic common diseases in women. According to the GLOBOCAN 2012 report, the age-standardized incidence of breast cancer and age-standardized mortality from this disease in China were 22.1 per 100000 and 5.4 per 100000, respectively [2]. Therefore, an efficient diagnosis and prevention method of breast diseases, especially breast adenosis, are urgently needed [3].

The IRT (infrared thermography) diagnosis is a relatively new functional detection technology that has been widely and increasingly applied to clinical diagnosis and research. By applying the IRT technology to the clinical diagnosis of various diseases, satisfactory results have been achieved [4-6]. Denoble developed a prediction method of the severity of knee arthritis, which is based on differences in mean temperature and standard deviation between the IRT images of knees, thereby verifying the objective reliability of IRT [7]. Kacmaz et al. proved that analyzing IRT images could be an effective method for pre-diagnosis of deep venous thrombosis [8]. Calin et al. demonstrated that IRT represented an effective tool for the diagnosis, evaluation, and monitoring of knee joint

lesions that could be used as a supplement to the standard methods of these types of lesions [9]. Elsa et al. conducted a comprehensive review of the application of medical IRT in the diagnosis of hand diseases based on results from 146 papers, which proved that the IRT could directly or indirectly provide useful diagnostic and monitoring information relating to the hand health state [10]. Ring and Ammer reviewed the application of IRT in the diagnosis of multiple diseases, including the inflammatory diseases, pain syndrome in complex regions, Reynolds disease and other diseases. They sucessfully identified the trend of using IRT in assessment of fever screening and some other areas of medicine by reviewing the related studies that had been published in recent years [11-13]. However, the analysis of a large number of IRT images is time-consuming and laborious, which increases the workload of doctors, and thus elongates the waiting time of patients. Therefore, the development of methods for fast analysis and classification of IRT images and more accurate disease prediction has become a focus of medical research.

With the advances of intelligent technologies, especially deep learning (e.g., multi-layer neural networks) in recent years, extensive research on these intelligent technologies has been conducted, and they have been successfully applied to the analysis of big biomedical data [14-16]. Some people used a neural network, genetic algorithms, and logistic regression to detect breast diseases and breast cancer [17-19]. Many scientists have applied the deep learning model to classify medical images and achieved very good results [20-22]. However, only a few studies have applied deep learning and IRT to the prediction and diagnosis of breast adenosis. Due to the good performance of deep learning in image recognition and its high recognition accuracy that is better than those of the classical artificial neural networks and other methods, here, we proposed a method for breast cancer detection based on IRT.

The proposed DCNN-LVQ network model combines the advantage of the good extraction capability of a DCNN (Deep convolutional neural network) with the high-accuracy classification capability of LVQ (learning vector quantization). The proposed network is trained with IRT images of human breast and then used to predict the breast adenosis. The proposed method can replace manual analysis of infrared thermography, saving both time and cost, providing an efficient and convenient diagnosis of human breast adenosis. The proposed DCNN-LVQ model is the first model applied to the IRT dataset.

Materials and methods

The DCNN model represents a deep feedforward artificial neural network that uses the weight sharing mechanism. The neurons in the first hidden layer are not connected with each neuron in the input layer. Therefore, compared with the traditional fully-connected feedforward network, the DCNN reduces the number of connections (i.e., the number of connection weights), thus reducing the number of network parameters, and achieving the information transmission via weight sharing. Compared with a conventional deep neural network, the DCNN can greatly reduce the number of network parameters and the computational complexity, representing a breakthrough regarding the image recognition success rate in the image processing field. However, the DCNN outputs the result via a Softmax function classifier (refer to Equation (1)), that is, the classification result of a sample is obtained by the probability-based classification. Therefore, when dealing with classification problems involving multiple features, especially when the sample size is small, the DCNN often fails to yield good classification results.

The LVQ algorithm represents a supervised learning algorithm that allows specifying the class to which the input is assigned. The LVQ algorithm can be trained to identify classes that consist of several unconnected regions. In addition, it is able to process a large amount of input data with a small amount of computation, which is highly desirable for processing the feature values extracted by the DCNN. One of the benefits of the LVQ algorithm is that it does not require to normalize and orthogonalize the input vector; instead, it calculates the distance between the input vector and the competition layer directly. Therefore, it denotes a simple, easy-to-implement, and highly robust algorithm.
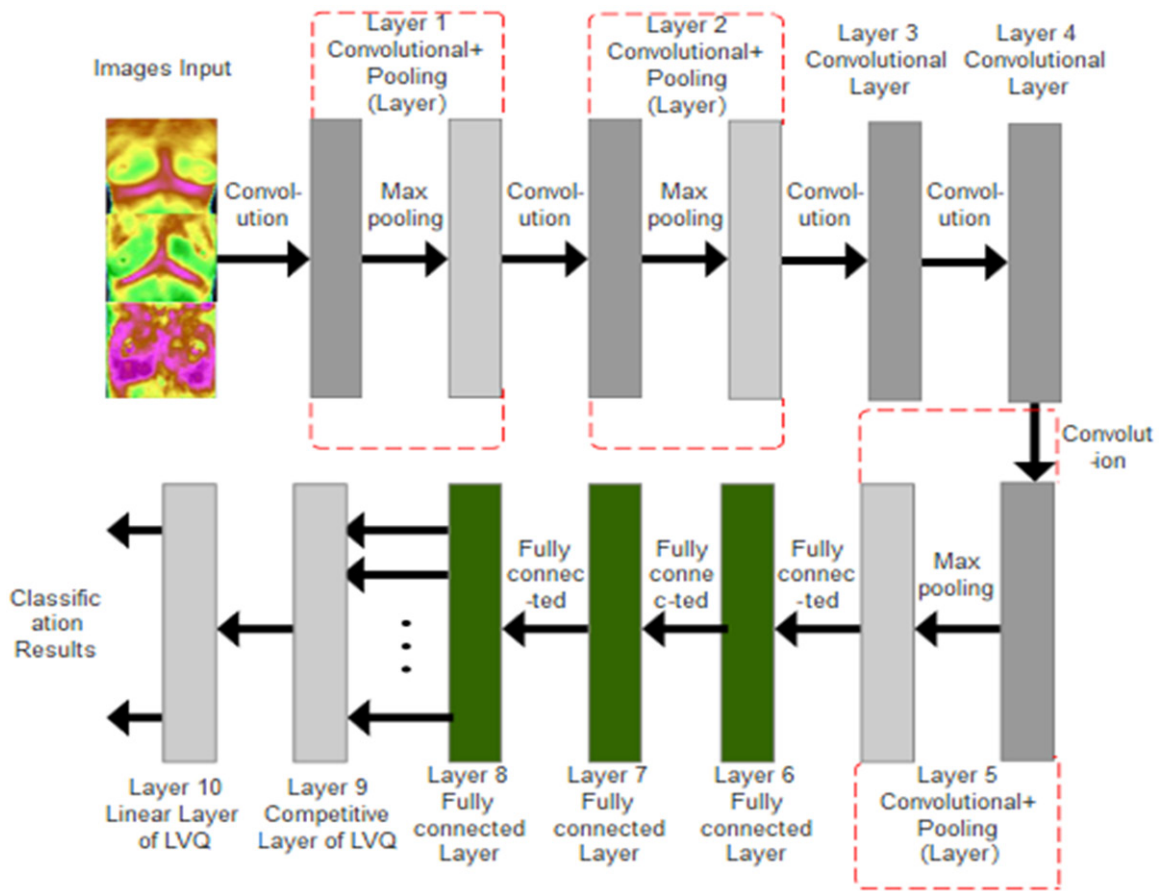
Figure 1. The structure of the proposed DCNN-LVQ classification model. The input image was passed to following convolutional-pooling layers and fully-connected layers in the form of a pixel matrix and then passed to the two layer of LVQ model.

$$y_k = \frac{exp(a_k)}{\sum_{i=1}^{n} exp(a_i)} \tag{1}$$

As mentioned above, this paper proposes a DCNN-LVQ network model that combines the advantages of the DCNN and the LVQ. The structure of the proposed model is shown in **Figure 1**. The proposed model consists of two parts. The first part is the DCNN that is used to extract the features of IRT images, and the second part is the LVQ model that is used to classify the features previously extracted by the DCNN.

Feature extraction by DCNN

The proposed DCNN-LVQ model uses DCNN as a feature extractor. The neural network structure was obtained by the analysis. The analysis results showed that network performance did not improve significantly with the number of layers, while network complexity raised greatly; finally, the 8-layer network structure was selected as the optimal one. The ReLU function given by Equation (2) was used as the activation function of both the convolutional layer and the fully-connected layer, and the loss function (the mean squared error function) was used to estimate the difference between the predicted value and the real value of sample. The first convolution layer consisted of 96 11×11 convolution filters, i.e., convolution kernels, with a step size of 4, and the corresponding pooling layer was the maximum pooling layer with a size of 3×3 and a step size of 2. The second convolutional layer consisted of 256 5×5 convolution filters with a step size of 1; the corresponding pooled layer was the maximum pooling layer with a size of 3×3 with a step size of 2. The third convolutional layer consisted of 384 3×3 convolutional filters with a step size of 1. The fourth convolutional layer consisted of 192 3×3 convolution filters with a

Table 1. Structural parameters of the neural network model

Number of layers	Convolution kernel size	Number of convolution kernels	Pooling layer mode	Pooling layer size
1	11×11	96	Max pooling	3×3
2	5×5	256	Max pooling	3×3
3	3×3	384	×	×
4	3×3	192	×	×
5	3×3	128	Max pooling	3×3
6			Fully connected	
7			Fully connected	
8			Fully connected	

step size of 1. The fifth convolutional layer consists of 128 3×3 convolution filters with a step size of 1, and the corresponding pooling layer was the maximum pooling layer with a size of 3×3 and a step size of 2. The sixth, seventh, and eighth layers were fully connected layers, the former two consisted of 4096 neurons each, and the later one had 1000 neurons. The structural parameters of the network are given in **Table 1**.

$$f(x)=\begin{cases} x & x \geq 0 \\ 0 & x \leq 0 \end{cases}. \qquad (2)$$

For the input layer, the input image was passed to the convolution layer in the form of a pixel matrix. Each convolution operation that was performed on the original image by the convolution kernel produced a new pixel, and different features were extracted when different convolution kernels performed the convolution operation.

The numbers of kernel rows and columns are usually set to odd numbers, which ensures that the two image sides are symmetrical after padding.

In the proposed network model, the convolutional layer reduces the loss of image information by adding zero at the edge of every pixel matrix. The output of the convolutional layer is passed to the pooling layer after being processed by the ReLU function.

The pooling layer is called either an undersampling or downsampling layer. It is mainly used for feature dimensionality and over-fitting reduction, compression of data and parameters, and improvement of fault tolerance of the model. In this paper, the pooling layer uses a

3×3 maximum value pooling process to compresses the output value of the convolution layer, thus obtaining a compressed and trimmed output value.

The last three layers are fully connected layers, and the output value of the last layer is used as the input to the LVQ classifier.

LVQ classifier

As shown in **Figure 2**, the LVQ network consists of two network layers. The first layer is the competition layer, and it classifies the input vectors. The second layer is the learning layer, and it converts the classification output of the competition layer into a different category.

In the LVQ neural network, there are 50 neurons in the competition layer (m=50), and 1000 neurons in the input layer (n=1000). These two layers are fully connected. Each neuron in the output layer is connected only to a group of neurons in the competition layer, and the corresponding connection weight is fixed and equal to one. During the training process, the weights between the input layer and the competition layer are gradually adjusted to the cluster center. After a sample is inputted into the LVQ network, some neurons in the competition layer win at the competition and become winning neurons following the learning rule of "winner is the king". The winning neuron's output is equal to one, and the outputs of the rest of the neurons are equal to zero. The output neurons that are connected to the group of winning neurons have the output equal to one, and the other output neurons have the output equal to zero; thus, the class pattern of the current input sample is obtained. The classes obtained by the learning process of the competition layer are taken as
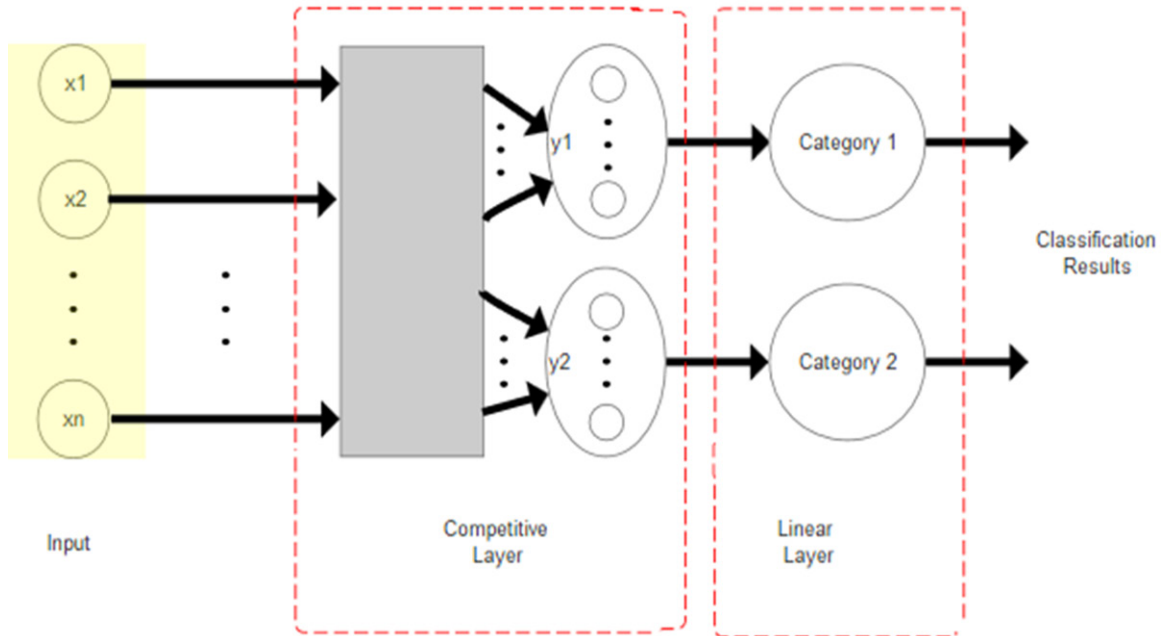
Figure 2. The structure of the LVQ neural network. The LVQ neural network consists of two network layers, the first layer is competition layer which classifies the input vectors, the second layer is learning layer which converts the classification output of competition layer into a different category.

subclasses, and the classes obtained by the learning process of the output layer are taken as target classes.

The learning process of the LVQ neural network combines the competitive learning rules and supervised learning methods, and the input and target vectors of the training sample appear in pairs. Thus, the training set can be expressed as:

$$\{(p_1,t_1),(p_2,t_2),...,(p_n,t_n)\} \tag{3}$$

where $t_i,(i=1,2,...,n)$ is a target vector of $m=50$ dimensions, corresponding to 50 neurons of the output layer; also, only one component is equal to one, while the remaining components are all equal to zero. Usually, each neuron of the competition layer is assigned to an output neuron, and the corresponding weight is set to one. In other words, the class of the output neuron category is specified, and it does not change during the training process. During the network training, only the weights of the input layer neurons to the competition layer neurons are changed [23].

DCNN-LVQ algorithm

The DCNN-LVQ algorithm includes the following steps.

(1) The weights of DCNN-LVQ are randomly initialized. The maximum number of iterations is denoted with t_m, the minimum error is denoted with ε, and the learning rate is denoted with α, and α>0.

(2) The image sample is fed to the DCNN input layer.

(3) The features extracted by the DCNN output layer are fed to the LVQ input layer as the input vector $X=(p_1,p_2,...,p_n)^T$.

(4) The distance between the neurons of the LVQ competing layer and the input vector is calculated by:

$$d_i = \sum_{j=1}^{n} (p_j - W_{ij}^1)^2. \tag{4}$$

The neuron with the smallest distance is taken as a winning neuron. The output of the winning neuron is equal to one, and the outputs of the remaining neurons are equal to zero.

(5) The output of the neurons connected to the group of winning neurons is equal to one, and those of the other output neurons are equal to zero; thus, the class pattern of the input sample is obtained. The class pattern resulted is then compared with the label t_i. If the network classification result is consistent with the label,

the weight is adjusted to be close to the input sample, which is expressed as:

$$W^1_{j*}(t + 1) = W^1_{j*}(t) + \eta(t)[X - W^1_{j*}(t)]. \qquad (5)$$

When the network classification result is inconsistent with the label, the weight is adjusted in the opposite direction to the input sample, which is expressed as:

$$W^1_{j*}(t + 1) = W^1_{j*}(t) - \eta(t)[X - W^1_{j*}(t)], \qquad (6)$$

where t denotes the number of iterations, W^1_{j*} denotes the weight of the connection from the input layer to the j th neuron of the competition layer, $\eta(t)$ denotes the learning rate of the t th iteration, and lastly, X denotes the input vector. The weights of other non-winning neurons remain unchanged [24].

(6) The learning rate is updated by:

$$\eta(t) = \eta(0)(1 - \frac{t}{t_m}), \qquad (7)$$

where t_m denotes the maximum number of iterations.

(7) The error function is calculated by:

$$E = \frac{1}{2}\sum_{i=1}^{P}\sum_{j}^{M}(x^i_j - a^i_j)^2, \qquad (8)$$

where x^i_j represents the target value (label value) of the j th neuron of the LVQ output layer when the input is the i th sample, a^i_j represents the actual output value of the j th neuron of the LVQ output layer when the input is the i th sample; P denotes the total number of samples, and M denotes the number of neurons in the output layer. If $E \le \varepsilon$ or $t = t_m$, the training terminates or; otherwise, the algorithm goes to Step (8) [23].

(8) Update the weights of the DCNN output layer neurons by the BP algorithm [24], which is given by:

$$\delta^{i,l}_j = -\frac{\partial E}{\partial a^{i,l}_j}f' = (x^i_j - a^i_j)f', \qquad (9)$$

$$W^l = W^l - \alpha\sum_{i=1}^{m}\delta^{i,l}(a^{i,l-1})^T, \qquad (10)$$

$$b^l = b^l - \alpha\sum_{i=1}^{m}\delta^{i,l}, \qquad (11)$$

where W denotes the weight matrix, b denotes the offset, α denotes the learning rate, δ is the gradient error, l represents the number of lay-

ers, i represents the i th image sample, j represents the j th neuron, a represents the neuron output value, T stands for transpose, f represents the neuron activation function. Since it is impossible to find the derivative of the ReLU function at the origin, here, the derivative of the ReLU function at the origin is set to zero.

(9) Update the weights of the fully-connected layer neurons in the DCNN hidden layer by the BP algorithm [24].

The weights and offsets of the network are updated according to Equations (10) and (11), respectively, and the gradient error is calculated by:

$$\delta^{i,l}_j = (\sum_{k=1}^{h}\delta^{i,l+1}_k w^{i,l+1}_k)f', \qquad (12)$$

where k denotes the k th neuron of the (l+1) th layer.

(10) Update the weights of the convolution kernel neurons in the DCNN convolutional layer [24] by:

$$W^l = W^l - \alpha\sum_{i=1}^{m}\delta^{i,l} * a^{i,l-1}, \qquad (13)$$

$$b^l = b^l - \alpha\sum_{i=1}^{m}\sum_{u,v}(\delta^{i,l})_{u,v}, \qquad (14)$$

where $\sum_{u,v}(\delta^{i,l})_{u,v}$ represents the process of obtaining the sum of the elements of each sub-matrix of $\delta^{i,l}$. After all the weights of the convolutional layer are updated, the algorithm goes to Step (2).

Experiment

Data source: In this study, a set of chest IRT images of patients chosen from the image repository of a certain hospital in Guangdong Province, China, were used. All data were arranged based on personal information of examinees, such as name, gender and age; name was hidden to protect the privacy of the examinees and data-using were agreed by examinees. In order to exclude the external influencing factors, the criteria for picking sample images were chosen, such that to ensure that the images were taken at the same ambient temperature and shooting distance and that all the patients were in a calm state at the time of imaging. Each sample was examined by the experts and was given a judgment on whether the person in question was a healthy
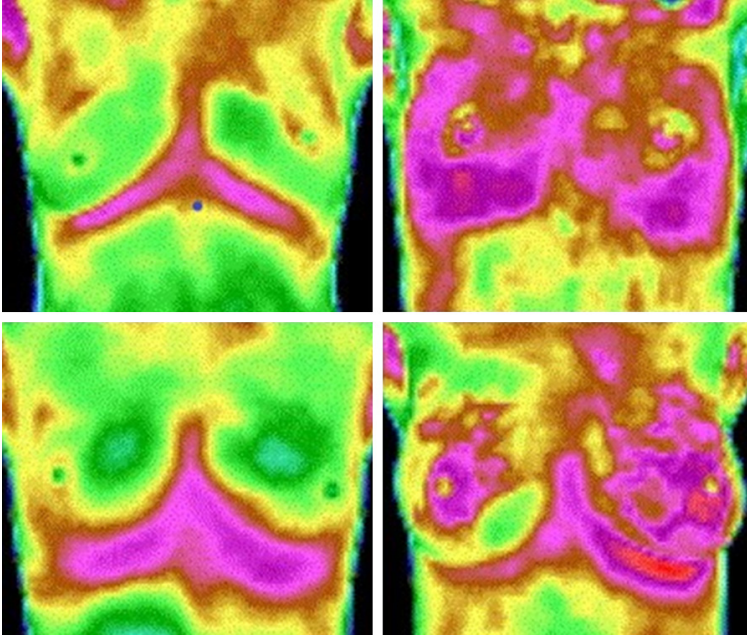
Figure 3. Images of negative (healthy person) and positive (patient of breast adenosis) samples. In an IRT image of the human body, different color gradations represent different temperature zones. Generally speaking, the IRT image of a healthy person shows a relatively even and stable temperature distribution. In contrast, the IRT image of a patient with breast adenosis is characterized by relatively uneven and unstable temperature distribution. Some of the samples used in this study are shown in **Figure 3**, where the left two images denote the negative samples (healthy person), and the right two denote positive samples (patient with breast adenosis).

person or a patient with breast adenosis, and accordingly labeled as a negative or positive sample, respectively. In this study, disease identification was conducted using 1000 images collected from year 2013 to 2019, among which 521 were positive samples, and 479 were negative ones. In an IRT image of the human body, different color gradations represent different temperature zones; namely, white represents a high-heat zone, red represents a hot zone, yellow represents a moderate-temperature zone, green represents a cool zone, blue represents a cold zone, and purple represents a super-cool zone. Thus, when the patient's chest has a lesion, the IRT image will show abnormal temperature distribution. Generally speaking, the IRT image of a healthy person shows a relatively even and stable temperature distribution. In contrast, the IRT image of a patient with breast adenosis is characterized by relatively uneven and unstable temperature distribution. Some of the samples used in this study are shown in **Figure 3**,

where the left two images denote the negative samples (healthy person), and the right two denote positive samples (patient with breast adenosis).

Data preprocessing: A total of 1,000 images (RGB) of breast area with a resolution of 150× 104 pixels were cropped from 1,000 larger images. As already mentioned, 521 of them were positive samples, and the remaining 479 were negative samples and these 1000 images are all chosen from patients in a certain hospital in Guangdong Province China from year 2013 to 2019. These 1000 breast images were used to create the general data set D, so that the numbers of positive and negative samples were roughly the same. The training set was denoted with Tr, which included 800 images chosen from year 2013 to 2018, and the test set was denoted as Te, which included the remaining 200 images from samples in 2019, so that Te is independent and irrelevant to Tr.

In order to distinguish the severity of patient's condition further, 521 positive samples were used as dataset *M*, and this dataset was divided into two sub-datasets, patients with a severe condition (327 images) and patients with a light condition (194 images). The division was conducted by experts in the field. The images of the patients with a severe condition were used as positive samples, and the images of the patients with a light condition were used as negative samples. The training set was denoted with Tr', which included 436 images chosen from year 2013 to 2018, and the test set was denoted as Te', which included the remaining 85 images from samples in 2019, so that Te' is independent and irrelevant to Tr'. An example of the samples is shown in **Figure 4**, where the left image represents the image of a patient with a light condition, and the right one is the image of a patient with a severe condition.
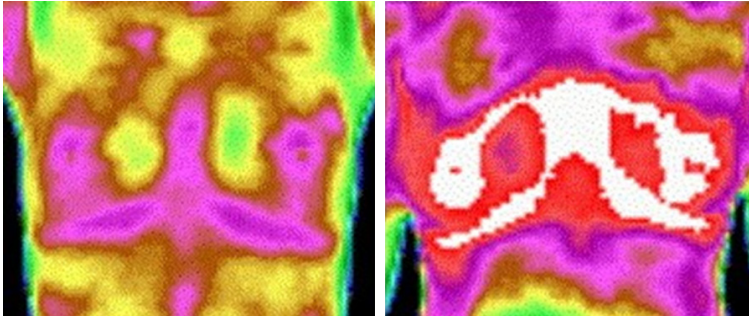
Figure 4. Images of patients with a light condition (left) and severe condition (right). In order to distinguish the severity of patient's condition further, positive samples (patient with breast adenosis) were used as dataset, and this dataset was divided into two sub-datasets, patients with a severe condition and patients with a light condition. The division was conducted by experts in the field. The images of the patients with a severe condition were used as positive samples, and the images of the patients with a light condition were used as negative samples.

These 1000 images included people over the age of 20, and details of dataset *D* and *M* are shown in **Tables 2** and **3**.

Operation: First, the experiment with the DCNN-LVQ model (refer to **Figure 1**) was carried out by using test set. The prediction accuracy of the experiment was taken as the prediction accuracy of the model. The prediction accuracy of experiment was calculated by:

$$a = \frac{N_c}{N_t}, \tag{15}$$

where N_c represented the number of correct predictions, and N_t represented the total number of samples in the test set.

The maximum number of loops was set to 1,000, the minimum error was set to 0.0001, and the initial learning rate was 0.8.

The results of experiments are shown in the following section.

Results

The infrared thermal imaging systems use photoelectric technology to detect infrared signals with a wavelength from a specified range, which are radiated from the object, calculate the temperature values of different spots, and convert the infrared signals into images and graphics visible to human eyes. In this section, the DCNN-LVQ model is used to predict the conditions of the human body based on the temperature distribution indicated by the infrared image.

The prediction accuracies of the DCNN-LVQ model obtained in the experiments using datasets *D* (refer to section *data preprocessing*) and D" (D mixed with noise) were 97.52% and 95.11% respectively; and the prediction accuracies of the experiments using datasets *M* (refer to section *data preprocessing*) and M" (M mixed with noise) were 94.66% and 92.17% respectively, which can be considered as satisfactory results. In the predictions based on datasets *D* and *M*, the loss function and training accuracy of the DCNN-LVQ are shown in **Figures 5** and **6**, respectively.

Model evaluation

In order to evaluate the performance of the DCNN-LVQ model further, it was compared with four related models, including the DCNN-LVQ, the AlexNet, the CNN (a convolutional neural network with two convolution layers, two pooling layers, and one fully connected layer), and the GoogleNet. In comparison, datasets D and D" (mixed with noise) were used. The prediction accuracies were calculated by Equations (15). The results obtained in the model comparison are given in **Table 4**.

Furthermore, the same models were compared on datasets M and M" (mixed with noise). The prediction accuracies were calculated by Equations (15). The comparison results are presented in **Table 5**.

As presented in **Tables 4** and **5**, the DCNN-LVQ outperformed all other models in prediction based on IRT images on all datasets, and it had the highest robustness.

This paper proposed a DCNN-LVQ prediction model based on a deep convolutional neural network and learning vector quantization algorithm. The image features are first extracted by the first eight layers of the DCNN and then classified by the LVQ algorithm. The proposed

Table 2. Details of dataset *D*

Details / Age group (proportion)	Proportion of climacteric syndrome (number of samples)	Proportion of obesity (number of samples)	Proportion of people with irregular diet (number of samples)	Proportion of people with other organ diseases (number of samples)	Proportion of endocrine disorders (number of samples)	Proportion of career people (number of samples)
Between 20 and 25 (10%)	0% (0)	8.0% (8)	45.0% (45)	1.0% (1)	0% (0)	30.0% (30)
Between 25 and 35 (25%)	0% (0)	8.8% (22)	54.8% (137)	4.8% (12)	10.0% (25)	95.2% (238)
Between 35 and 45 (27%)	4.1% (11)	11.1% (30)	21.9% (59)	17.8% (48)	30.0% (81)	97.8% (264)
Between 45 and 55 (20%)	95.0% (190)	14.0% (28)	10.0% (20)	45.0% (90)	65.0% (130)	82.5% (165)
Over 55 (18%)	6.1% (11)	12.8% (23)	1.7% (3)	68.9% (124)	22.2% (40)	0% (0)

Table 3. Details of dataset *M*

| Details / Age group (proportion)	Proportion of climacteric syndrome (number of samples)	Proportion of obesity (number of samples)	Proportion of people with irregular diet (number of samples)	Proportion of people with other organ diseases (number of samples)	Proportion of endocrine disorders (number of samples)	Proportion of career people (number of samples)
Between 20 and 25 (0%)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)
Between 25 and 35 (15%)	0% (0)	19.2% (15)	84.6% (66)	9.0% (7)	23.1% (18)	100% (78)
Between 35 and 45 (47%)	3.3% (8)	10.2% (25)	20.4% (50)	18.4% (45)	29.4% (72)	99.6% (244)
Between 45 and 55 (25%)	100% (130)	16.2% (21)	15.4% (20)	57.7% (75)	88.5% (115)	96.2% (125)
Over 55 (13%)	13.2% (9)	25.0% (17)	2.9% (2)	98.5% (67)	57.4% (39)	0% (0)
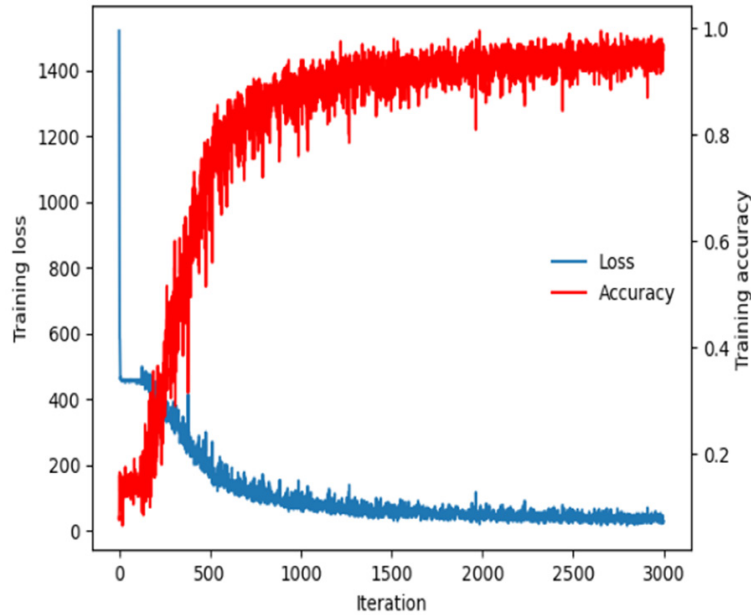
Figure 5. The loss function and training accuracy of the DCNN-LVQ model on dataset Tr of D. The loss function (the mean squared error function) is used to estimate the difference between the predicted value and the real value of sample, which is calculated by equation (8). The training accuracy is calculated by equation (15), which shows that the performance of the model on dataset Tr of D.
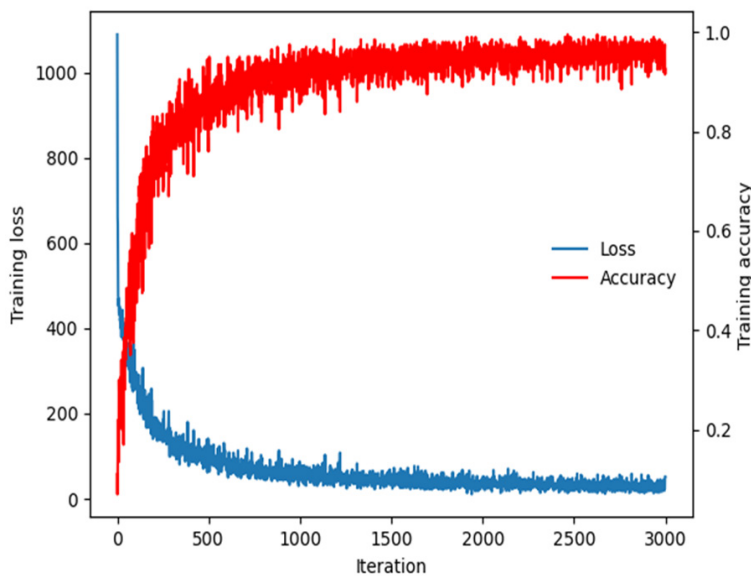

Figure 6. The loss function and training accuracy of the DCNN-LVQ model on dataset Tr' of M. The loss function (the mean squared error function) is used to estimate the difference between the predicted value and the real value of sample, which is calculated by equation (8). The training accuracy is calculated by equation (15), which shows that the performance of the model on dataset Tr' of M.

mental datasets, the DCNN-LVQ achieved the accuracies of 97.52% and 94.66% in the prediction of breast adenosis based on IRT images. Besides, the proposed model was compared with three common prediction models, and the comparison results showed that the proposed model overperformed the other models in IRT image recognition. Thus, the proposed model can provide breast adenosis prediction with high accuracy. In addition, the proposed model is highly robust; therefore, it has a high potential to be applied to other IRT image recognition problems.

Discussion

IRT, a non-invasive, non-contact, and 2D imaging thermal measurement technology, has many advantages over the traditional body temperature measurement technologies. It has been increasingly applied to medical research, disease diagnosis, and health assessment in recent years. When a tumor is very small, even the imaging technology with high resolution will find it difficult to identify it. However, unlike other imaging technologies, IRT may recognize the temperature difference of the tumor, although it may not determine the shape of a small tumor. Then, the doctors are able to diagnose and treat the disease of patients based on their abnormal temperature, such that preventive diagnosis can be made in the early stage of the disease. IRT has been successfully applied to the diagnosis of breast tumors, diabetic neuropathy, peripheral vascular diseases, and other diseases. It has also been widely used in the detection and analysis in the fields of sports

model is verified by the experiments. The experimental results showed that on two experi-

Table 4. The disease prediction accuracies of different prediction models

Dataset	Parameter	Prediction model			
		DCNN-LVQ	AlexNet	CNN	GoogleNet
D	Accuracy	0.9752	0.9231	0.8632	0.9379
D"	Accuracy	0.9511	0.9032	0.8136	0.9024

Table 5. The disease prediction accuracies of different models

Dataset	Parameter	Prediction model			
		DCNN-LVQ	AlexNet	CNN	GoogleNet
M	Accuracy	0.9466	0.9013	0.8875	0.9189
M"	Accuracy	0.9217	0.8832	0.8536	0.8973

medicine, obstetrics and gynecology, dermatology, cardiovascular department, pediatric physiology, and other diseases.

As a convolutional layer has the parameter sharing mechanism, while a pooling layer can reduce calculation redundancy, the DCNN can therefore achieve a satisfactory image recognition efficiency. The convolutional and pooling layers of the DCNN possess the features of automatic extraction of image, thus avoiding the time-consuming manual feature extraction. Still, better classification results can be achieved by further improvement of the proposed DCNN-LVQ model. The LVQ algorithm can be trained to identify a class consisting of several unconnected areas. It can process a large amount of input data with a small amount of calculation, and does not require normalization and orthogonalization of the input vector. All in all, the LVQ algorithm denotes an easy-to-implement and robust algorithm. The proposed DCNN-LVQ network model combines the feature of extraction capability of the DCNN with the LVQ capability of highly-accurate classification, which makes it able to achieve a satisfactory classification accuracy of IRT images.

The method of adding zeros at the edge, which is used in this study, represents an effective technique for improving image recognition performance. Namely, it can reduce the loss of image information caused by convolution operation, and can yield good edge extraction result.

The proposed model represents an improved AlexNet model. The experimental results show-ed that the prediction accuracy of the proposed model was higher than that of the AlexNet model. Under the same conditions, including the software and hardware environments, the running time of the proposed model was not much shorter than that of AlexNet with eight hidden layers. The proposed model could also realize the real-time prediction of the infrared thermal image. In addition, due to the fault tolerance and robustness of the neural network, the DCNN-LVQ network model has a strong universality, so it can be widely used in other areas of image recognition.

As the IRT systems measure only the body surface temperature, the result is affected by a variety of factors. The body surface temperature of a person may change considerably under some special conditions, such as skin dysfunction and breastfeeding (for some women), resulting in an incorrect recognition result. Although in the experiments, the judgment on whether an image was a positive or negative sample had been given by the experts in the field, misjudgments were unavoidable. In addition, the preliminary analysis of the IRT images by the experts was very time-consuming. However, we assume that positive and negative samples can be accurately identified by an autonomous machine using an unsupervised algorithm.

There are many unsupervised learning algorithms, such as the clustering algorithm. The clustering algorithm represents a type of un-labeled algorithm, which classifies samples based on certain rules, such as Euclidean distance between samples. Due to the similarity between normal (or pathological) breast images, in the absence of a label value, the clustering method can be used for future image classification.

Disclosure of conflict of interest

None.

Address correspondence to: Longlong Liu, School of Mathematical Sciences at Ocean University of China, Qingdao 266000, PR China. Tel: +86-13173237966; Fax: +86-0532-85901752; E-mail: liulonglong@ouc.edu.cn

References

[1]   Chen W, Liu XM and Yu DY. Diagnostic value of breast X-ray imaging report and data system for benign and malignant breast masses. Shanxi Medical Journal 2016; 45: 37-39.

[2]   Wang L, Zhang Y, Shi JF and Dai M. Analysis of disease burden of female breast cancer in China. Chinese Journal of Epidemiology 2016; 37: 970-976.

[3]   Zheng Y, Wu CX and Wu F. Current situation and development trend of female breast cancer mortality in China. Chinese Journal of Preventive Medicine 2011; 45: 150-154.

[4]   Jimenez-Pavon D, Corral-Perez J, Sánchez-Infantes D, Villarroya F, Ruiz JR and Martinez-Tellez B. Infrared thermography for estimating supraclavicular skin temperature and BAT activity in humans: a systematic review. Obesity 2019; 27: 1932-1949.

[5]   Jagadev P and Giri LI. Non-contact monitoring of human respiration using infrared thermography and machine learning. Infrared Phys Technol 104: 103-117.

[6]   Hillen B, Pfirrmann D, Ngele M and Simon P. Infrared thermography in exercise physiology: the dawning of exercise radiomics. Sports Med 2020; 50: 263-282.

[7]   Denoble AE, Hall N, Pieper CF and Kraus VB. Patellar skin surface temperature by thermography reflects knee osteoarthritis severity. Clin Med Insights Arthritis Musculoskelet Disord 2010; 3: 69-75.

[8]   Kacmaz S, Ercelebi E, Zengin S and Cindoruk S. The use of infrared thermal imaging in the diagnosis of deep vein thrombosis. Infrared Phys Technol 2017; 86: 120-129.

[9]   Brailescu, Monica C, Mologhianu, Gilda, Calin, Antonina M, Savastru, Roxana and Romeo M. A review of the effectiveness of thermal infrared imaging in the diagnosis and monitoring of knee diseases. Infrared Phys Technol 2015; 69: 19-25.

[10] Sousa E, Vardasca R, Teixeira S, Seixas A and Costa-Ferreira A. A review on the application of medical infrared thermal imaging in hands. Infrared Phys Technol 2017; 85: 315-323.

[11] Ring EF and Ammer K. Infrared thermal imaging in medicine. Physiol Meas 2012; 33: R33.

[12] Zhang L, Small GW, Haka AS, Kidder LH and Lewis EN. Classification of fourier transform infrared microscopic imaging data of human breast cells by cluster analysis and artificial neural networks. Appl Spectrosc 2003; 57: 14-22.

[13] He B, Guan Y and Dai R. Classifying medical relations in clinical text via convolutional neural networks. Artif Intell Med 2018; 93: 43-49.

[14] Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB and Liang J. Convolutional neural networks for medical image analysis: full training or fine tuning? IEEE Trans Med Imaging 2016; 35: 1299-1312.

[15] Milletari F, Navab N and Ahmadi SA. V-net: fully convolutional neural networks for volumetric medical image segmentation. IEEE Fourth International Conference on 3D Vision (3DV) - Stanford, CA, USA 2016; 32: 33-45.

[16] Ahmad FK, Deris S and Othman NH. The inference of breast cancer metastasis through gene regulatory networks. J Biomed Inform 2012; 45: 350-62.

[17] Ahmad F, Isa NA and Hussain Z. A genetic algorithm-based multi-objective optimization of an artificial neural network classifier for breast cancer diagnosis. Neural Comput Appl 2013; 23: 1427-1435.

[18] Faradmal J, Soltanian AR, Roshanaei G, Khodabakhshi R and Kasaeian A. Comparison of the performance of log-logistic regression and artificial neural networks for predicting breast cancer relapse. Asian Pac J Cancer Prev 2014; 15: 5883-8.

[19] Sadoughi, Farahnaz, Ghaderzadeh, Mustafa, Fein, Rebecca, Standring and Arran. Comparison of back propagation neural network and back propagation neural network based particle swarm intelligence in diagnostic breast cancer. Applied Medical Informatics 2014; 34: 22-30.

[20] Harbeck N, Kates R, Ulm K, Graeff H and Schmitt M. Neural network analysis of follow-up data in primary breast cancer. Int J Biol Markers 2000; 15: 116-22.

[21] Ecker BL, Lee JY, Sterner CJ, Solomon AC, Pant DK, Shen F, Peraza J, Vaught L, Mahendra S, Belka GK, Pan TC, Schmitz KH and Chodosh LA. Impact of obesity on breast cancer recurrence and minimal residual disease. Breast Cancer Res 2019; 21: 41.

[22] Lisa Y, Meng YL, Ya QL, Feng Y and Jing H. Palm vein classification based on deep neural network and random forest. Progress in laser and Optoelectronics 2019; 56: 10-19.

[23] Zhang X, An F, Chen L and Mattausch HJ. Reconfigurable VLSI implementation for learning vector quantization with on-chip learning circuit. Jpn J Appl Phys 2016; 55: 04-12.

[24] Ben-Nun T and Hoefler T. Demystifying parallel and distributed deep learning: an in-depth concurrency analysis. ACM Comput Surv 2018; 52: 4-16.