

Original Article

Identification of human diseases of the lower limbs based on medical infrared imaging and a CNN-LM model

Zichen Lu^{1*}, Di Ma^{2*}, Longlong Liu²

¹Cloud Computing Center, Chinese Academy of Science, Dongguan 523000, P. R. China; ²School of Mathematical Sciences, Ocean University of China, Qingdao 266000, P. R. China. *Equal contributors.

Received November 18, 2019; Accepted December 22, 2019; Epub March 15, 2020; Published March 30, 2020

Abstract: This paper seeks a stable and efficient method to analyze and identify medical infrared thermograms of human diseases in the lower limbs and provides an accurate auxiliary method for the diagnosis of these diseases. We propose a CNN-LM model, that is, the Levenberg-Marquard (LM) algorithm which is used to update the weight of the hidden layer in the convolutional neural network (CNN), and the feature extraction function of the convolutional neural network is combined with the simple and efficient LM algorithm to construct a new image recognition model. Using real medical infrared thermograms from the legs of patients in a Chinese hospital as test samples, using the initial expert medical diagnostic as the label, the model was trained and tested using the seven-fold cross-validation. The recognition accuracies in the original sample set and noise-added sample set were 96.72% and 95.90%, respectively; both of which are significantly better than other methods and achieved satisfactory results.

Keywords: Convolutional neural network, LM algorithm, medical infrared imaging

Introduction

Scientific and technological developments continue to benefit the medical field. Presently, medical infrared imaging is a convenient and effective auxiliary means for detecting diseases [1-3]. The temperature distribution of a normal human body is rather stable and has certain characteristics. When a disease state or functional changes in a certain part of the human body are present, blood flow changes accordingly, resulting in a change in the local temperature. Thus, a thermogram of the human body is collected using the medical infrared imaging system. After a medical professional analyzes the thermogram, the location of the lesion, the nature of the disease, and the degree of pathology can be judged, and an initial determination of the patient's condition can be made [4]. The literature shows that diseases of the human leg cannot be underestimated, as lower extremity venous thrombosis, lower extremity tumors, skin cancer, calf arterial injury, rheumatoid arthritis, and other such diseases

all pose risks to life and health. However, current medical approaches are incomplete and cannot always discover the disease in a timely manner [5, 6]. At the same time, approximately 80% of cancer patients are only discovered and correctly diagnosed in the middle and late stages of the disease [7, 8]. In addition, having to visually assess a large number of thermograms can lead to fatigue and possible misdiagnosis on the part of the medical professional [9]. This issue in the medical industry indicates that it is extremely important to find an efficient and accurate method for disease classification and identification [10]. In order to solve the above problems, this paper uses CNN's powerful feature extraction and autonomous learning ability, and uses the LM algorithm-optimized CNN model rather than the traditional manual method in the analysis and identification of the patient's leg thermogram. This can not only greatly reduce the diagnosis time, but can also predict early conditions and help patients avoid missing out on the best treatment window for their condition.

With the development of deep learning, convolutional neural networks are widely used in various fields due to their powerful pattern recognition capabilities [11, 12]. Liu Yulong et al. applied artificial neural networks and deep learning techniques to the identification and prediction of biological information and achieved excellent results [13-16]. In the medical field, Maria Baldeon-Calisto used CNN to complete the segmentation of medical images [17]; Måns Larsson et al. proposed a fully automated abdominal organ segmentation system using CNN, and achieved a DICE coefficient of 0.767 [18]; Zhehuan Zhao et al. constructed a new dual-phase multi-mode automatic diagnosis system for brain tumors, which can accurately locate and detect brain tumor types [19]; Dalal Bardou proposed an automatic pattern recognition system based on CNN and improved upon the limitations of traditional auscultation technology, resulting in a recognition accuracy rate of up to 95.56% [20]. Aykanat M et al. also confirmed the accuracy and feasibility of CNN classification of lung sounds [21].

However, during the back propagation of CNN error, the traditional gradient descent algorithm has been criticized for its low training speed. Although some researchers have improved the gradient descent algorithm, there are still limitations [22, 23]. Therefore, selecting a suitable weight updating algorithm can effectively improve the performance of CNN, making it an urgent problem to be solved. Based on the advantages of CNN in feature extraction and the wide adaptability and speed of LM algorithm in pattern recognition, this paper combines these two methods and proposes a CNN model based on the LM algorithm and applies it in the classification and identification of human diseases of the lower limbs. The experimental results show that the proposed algorithm is superior to the traditional algorithm in classification accuracy and speed, and it is an effective algorithm for medical infrared imaging-based recognition of human lower limb diseases.

Methods

Training convolutional neural networks can be divided into two parts: the first stage is the forward propagation stage, that is, the stage where the data is transmitted from the lower layer to the upper layer; the second stage is the

back propagation stage, that is, when the results from forward propagation are inconsistent with expected results, the discrepancy is propagated from the upper layer to the lower layer in an updated stage. This paper mainly optimizes the algorithm at this stage.

Forward propagation

In forward propagation, the input graphic data is subjected to convolution and pool processing, and the feature vector is proposed. The feature vector is transmitted into the fully connected layer, and the result of classification recognition is obtained in the output layer. When the output result is similar to the expected value, the result is output; otherwise, error correction is performed by backpropagation. The basic structure of the network is shown in **Figure 1**.

In the input layer, the input X to the network is usually an image, and the pixel values of the image form a matrix. In the convolutional layer, the input samples are convolved with a learnable convolution kernel. Convolution of different convolution kernels can extract different features, and the feature maps of these different features are used together as input data of the pooling layer. The j th feature map output in the convolutional layer a_j is:

$$a_j = f\left(b_j + \sum_{i \in M_i} x_i * k_{ij}\right) \quad (1)$$

Where, $f(\cdot)$ represents the activation function, b_j is the j th offset, x_i represents the i th pixel value of the input layer, $*$ represents the convolution operation, k_{ij} represents the convolution kernel, and M_i represents a subset of the input feature map. The activation function $f(\cdot)$ of the convolutional layer maps the convolved kernel convolved values to values within a specific boundary. Common activation functions are Sigmoid, Tanh, ReLU, etc. In the pooling layer, the feature map of the upper layer is subjected to down sampling processing, and a plurality of pixel values are compressed into one for feature extraction to form a new feature map. The j th feature map output b_j in the pooled layer is:

$$b_j = \text{pooling}(a_j) \quad (2)$$

Where *pooling* represents a pooling operation, which divides the feature map a_j into a number of rectangular image blocks of the same size

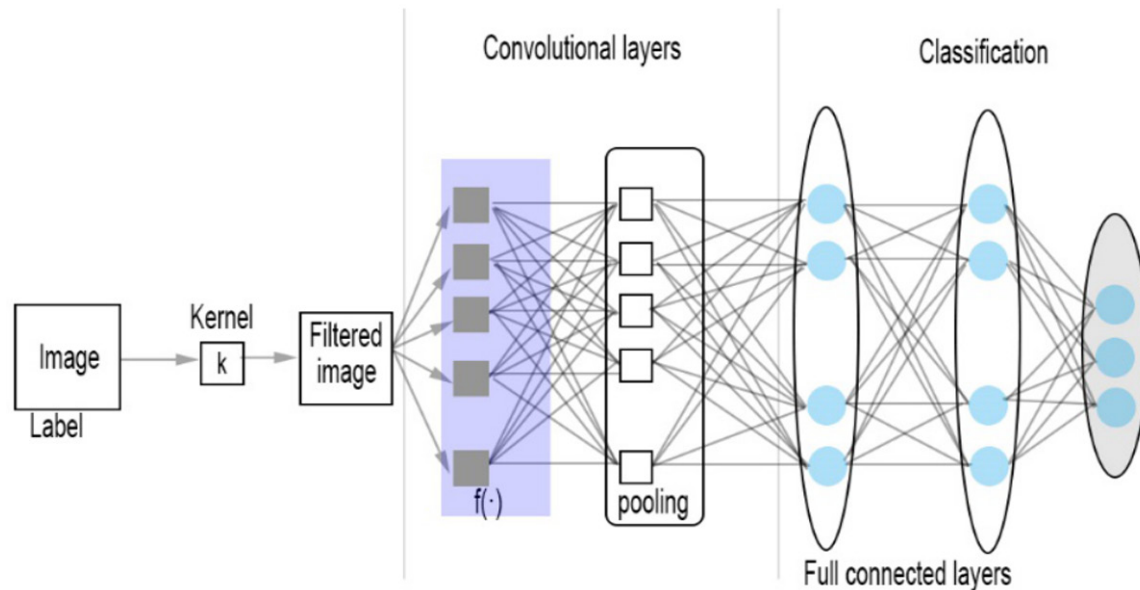


Figure 1. CNN structural diagram.

that do not overlap, and operates on the pixels in each image block. Common pooling operations include maximum pooling and average pooling.

The pooled feature map is transformed into a list of nodes to form a feature vector Y . After fully connecting the layers, the actual output is finally calculated as:

$$O = f(W^T \cdot Y + b) \quad (3)$$

Where W is a dimension vector representing the weight of the hidden layer; b is the offset.

Backpropagation

Backpropagation serves to adjust the error of the output layer when adjusting the network weight, and then gradually updates the weight of the hidden layer according to the weight update algorithm. Common weight update algorithms include Newton's method, Gauss-Newton method, steepest descent method, and elastic algorithm. However, for large networks, they all have corresponding limitations: Newton's method calculates the Hessian matrix at each step and its speed is slow. The Gauss-Newton method uses the Jacobian matrix to approximate the Hessian matrix, which improves the efficiency of the algorithm, but when the Hessian matrix is not full rank, it cannot be iterative; the steepest descent method

is fast when applied to pattern recognition, but the performance will deteriorate as error decreases; the elastic algorithm is more practical in moderately sized networks, and has relatively low storage requirements, but it does not perform well in large networks.

The LM algorithm can be used to solve nonlinear least squares problems and is often used in curve fitting and other occasions [24-27]. The key to the realization of the LM algorithm is to use the model function to perform a linear approximation of the estimated parameter vector in its neighboring region, ignoring derivative terms above the second order, and thus transforming into a linear least squares problem, since there is no need to calculate and store a large Jacobian matrix, it has better comprehensive performance in large networks, possesses faster convergence speed and higher precision, and can achieve smaller mean square error than other algorithms.

In this paper, the LM algorithm is used to update the weights. The parameters of the convolutional layer and the fully connected layer are generated by random uniform initialization. The specific process is as follows:

For the i th sample, the actual output value o_i is compared to the expected label value y_i to obtain the error e_i of the training sample.

$$e_i = o_i - y_i = \begin{pmatrix} e_{1,i} \\ \vdots \\ e_{l,i} \end{pmatrix} \quad (4)$$

In the formula, $e_{k,i}$ represents the k th dimension of e_i and l represents the output vector dimension.

The total error E of the model is expressed as:

$$E = \frac{1}{2} \begin{pmatrix} e_{1,1}^2, \dots, e_{l,1}^2 \\ \vdots \\ e_{1,N}^2, \dots, e_{l,N}^2 \end{pmatrix} \quad (5)$$

Where, N is the total number of samples.

Assuming that the current step is from the k th step to the $k+1$ th step, the size of the hidden layer weight matrix is $n \times l$ dimension, in which W_k represents the current hidden layer weight matrix, W_{k+1} represents the weight matrix after updating once, and the updated weight formula is:

$$W_{k+1} = W_k - [J^T(W_k)J(W_k) + \mu_k I]^{-1} J^T(W_k) E_k \quad (6)$$

Where, μ_k is the adjustment factor used to control the iteration of the LM algorithm, here taken as 0.01; E_k represents the error after the k th update; I is the n th order unit matrix; $J(W_k)$ is the Jacobian matrix which is the differential of the current weight W_k , noted as $W^T = (w_1 \dots w_n)$, and the Jacobian matrix $J(W)$ is:

$$J(W) = \begin{bmatrix} \frac{\partial e_1}{\partial w_1} & \dots & \frac{\partial e_1}{\partial w_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial e_N}{\partial w_1} & \dots & \frac{\partial e_N}{\partial w_n} \end{bmatrix} \quad (7)$$

When the 1-norm of the error satisfies the given error limit $\|E_k\|_1 \leq \epsilon$ or reaches the highest number of iterations, the training is terminated and the CNN structure constructed in this paper is obtained.

Results

First, the network model constructed in this paper was trained using the medical infrared thermogram labeled as diseased, then the

trained model was used to predict disease in the test image set. The overall process flow is shown in **Figure 2**.

Data source

The 502 samples selected in this paper were all selected from medical thermograms provided by a Chinese medicine hospital in Guangdong Province. The use and publication of the images have been approved by the involved patients and hospitals. During thermogram capture, the medical infrared thermal imager is used to eliminate interference from external factors, that is, sampling at the same temperature, the same imaging distance, and taken while the body is at rest. Sample labeling is performed following a series of examinations and via the doctor's diagnosis, which approximates the real situation of the patient. The samples include: the first type of sample (no obvious disease in the lower limbs or partial abnormality that can heal on its own), the second type of sample (some body parts exhibit abnormal characteristics, possibility of disease, such as knee joint tuberculosis, calf arterial injury, erythema limb pain, rheumatoid arthritis, patella softening, etc., further examination is recommended); the third type of samples (obvious abnormalities are exhibited, the possibility of related diseases is greater, such as venous thrombosis, lower extremity tumors, skin cancer, etc., timely diagnosis is recommended to avoid worsening of the disease).

Pretreatment

Data screening: The existence of irrelevant data can likely cause over-fitting of the model, resulting in weaker generalizability of the model. Therefore, it is necessary to screen the data before the experiment. In this paper, the data is filtered manually, and 12 samples containing irrelevant data and damaged images are eliminated, and 490 samples are obtained at the end.

Unified resolution: In order to eliminate the influence of other factors on our reported results, image processing software was used to adjust the resolution of the original image, and images were uniformly changed to a resolution of $204 \times 304 \times 3$.

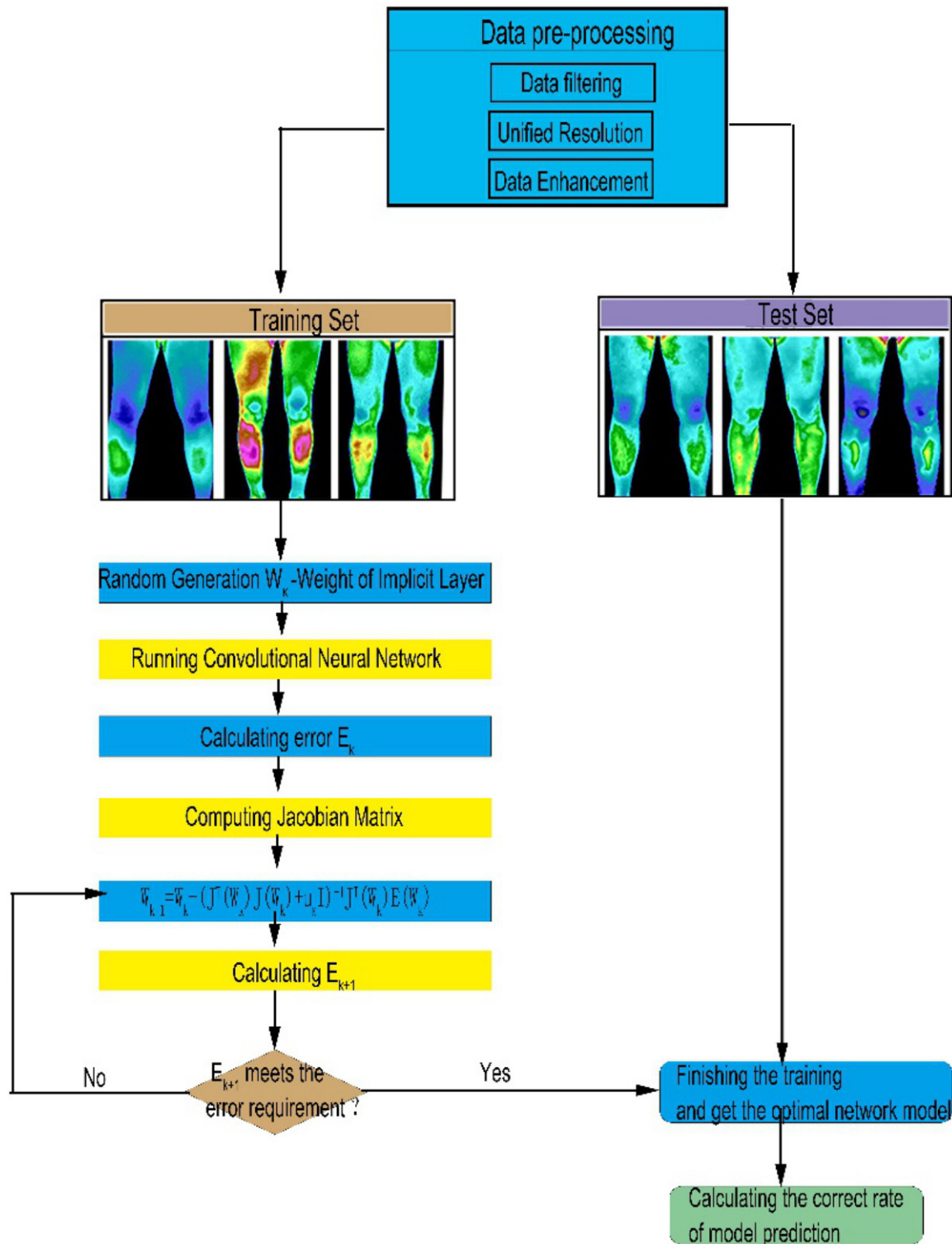


Figure 2. A flow chart illustrating the method presented in this paper.

Data enhancement: Data enhancement is a means by which more data can be obtained and minor changes can be made to existing data sets to increase the data set. Data

enhancement can prevent neural networks from learning unrelated information and can fundamentally improve overall performance. Due to the left-right symmetry of the human

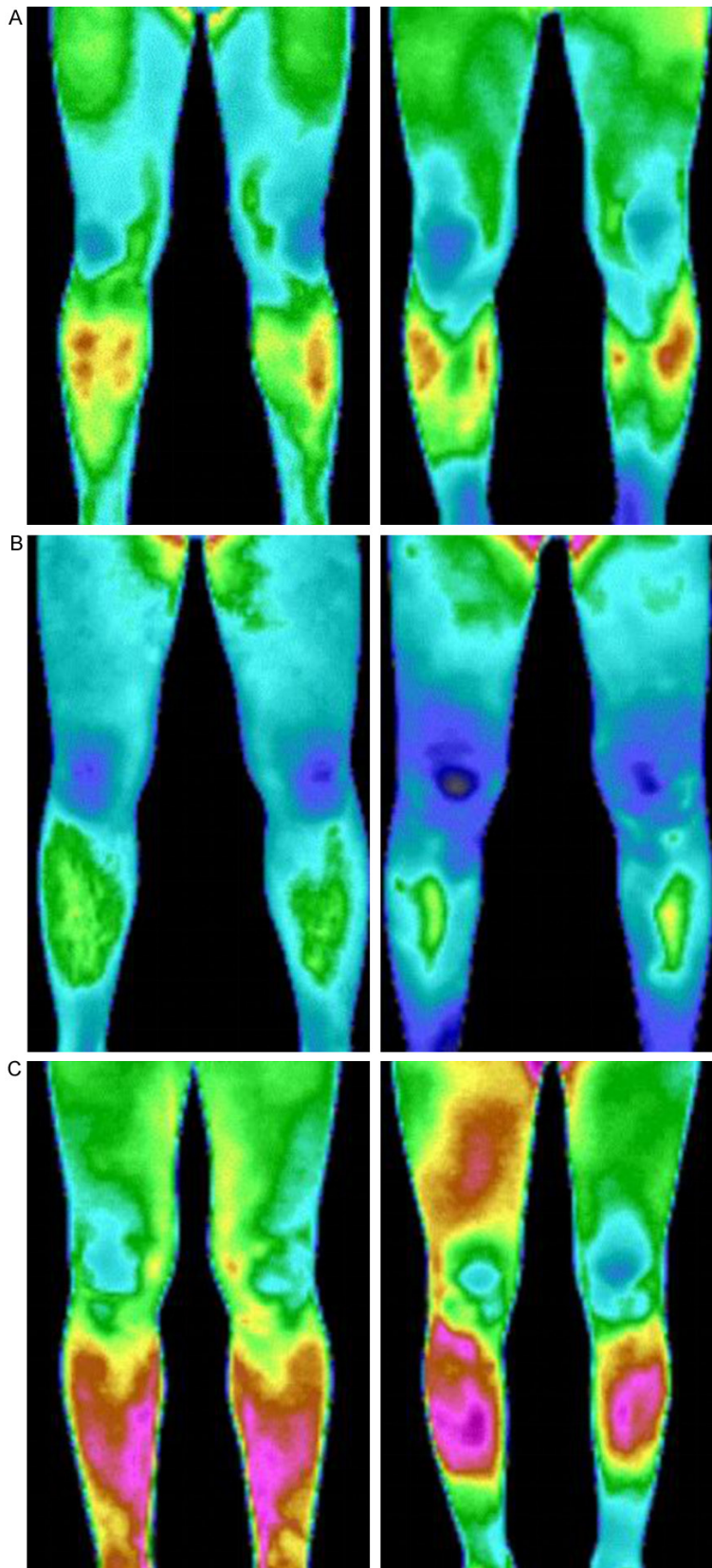


Figure 3. A. The first type of sample. B. The second type of sample. C. The third type of sample.

leg, the data set was enhanced by flipping the original picture horizontally.

After the data was preprocessed, a total of 980 samples were obtained, which constituted the data set D_1 , of these, 328 samples were of the first type, 328 samples were of the second type, and 324 samples were of the third type. Some samples are shown in **Figure 3**.

Experimental results

For the sake of simplicity, this paper builds a single-volume, single-pool convolutional neural network. The setting of the parameters is assisted by computer software and determined using the “trial and error method”. When a certain set of parameters can minimize network error, they were selected as the best set of parameters. There are 20 convolution kernels with a size of 5×5 in the reported network structure. The activation function of the convolutional layer is ReLU function. The average pooling layer size is 2×2 , the number of hidden layer nodes is 100, and the activation function is Sigmoidal. The size of the input layer is $204 \times 304 \times 3$, the weight initialization method is random initialization, the error limit is 0.03, and the maximum training times is set to 1000.

Using Matlab 2016a, the memory is 8.00GB, the CPU is Intel(R)4 Core(TM) i5-7200U@ 2.50GHz, and the GPU is NVIDIA GeForce 940MX. The neural network was trained on this computer and tested.

In the experiment, a seven-fold cross-check is used, that is,

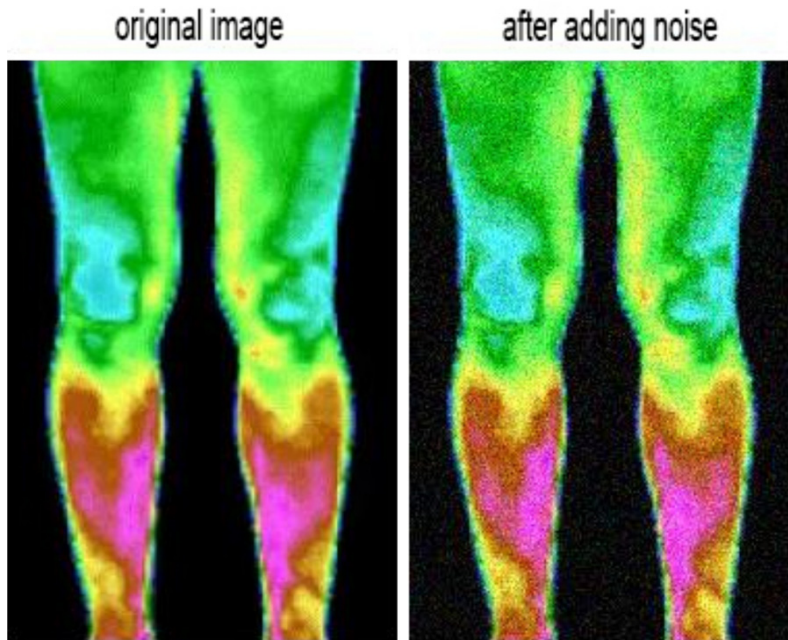


Figure 4. Comparison before and after noise addition.

858 samples are randomly selected as a training set to complete the training of the model. Then the remaining 122 samples are used as the test set, and the average accuracy over seven attempts is taken to determine the recognition accuracy a of the model.

$$a = \frac{1}{7} \sum \frac{N_c}{N_t} \quad (8)$$

In the formula, N_c represents the predicted correct quantity and N_t represents the total number of test sets.

The experimental results show that the recognition accuracy of the model is 95.90%.

Model evaluation

The above experiments show that the method can make a high accuracy judgment on the disease condition of the lower limbs in humans, can be used to perform a pre-diagnosis and generate early warning for related diseases, greatly reducing the time required before the patient receives treatment, and effectively prevent the worsening of disease.

To further verify the performance of the reported model, as shown in **Figure 4**, 400 original

pictures from D_1 are randomly selected, and noise with a mean value of 0 and a variance of 0.01 was added to form a new data set D_2 .

Based on the data set, the performance of four different algorithms used to adjust weight was compared. The results show that in both data sets, the model proposed in this paper has a comparative advantage in both accuracy and training speed.

To further explain the convergence speed of the LM algorithm, based on the D_1 , D_2 data sets, the iterative times of the gradient descent algorithm and the LM

algorithm are compared with the prediction accuracy, and shown in **Figures 5** and **6**.

As shown in the Figure, in the experiment, CNN-LM showed a clear advantage over the traditional algorithm, and the convergence speed increased significantly.

Discussion

Compared with other models, the CNN-LM model has a comparative advantage in recognition accuracy. In addition, the comparison of the analysis results before and after adding noise in Section 3.4 shows that the model in this paper is robust, specifically, even if the temperature of a certain part of the patient's leg is "artificially high" due to differences in individual constitution or external trauma, the prediction accuracy of the model does not change significantly, which shows that the results have higher credibility and can be generalized.

As observed in equation (6), if the proportional coefficient $\mu = 0$, then it is a Gauss-Newton method; if the μ value is very large, the LM algorithm approaches the gradient descent method, and each iteration is one step of success, i.e. μ is slightly reduced, so that as it approaches the error target, it gradually becomes similar

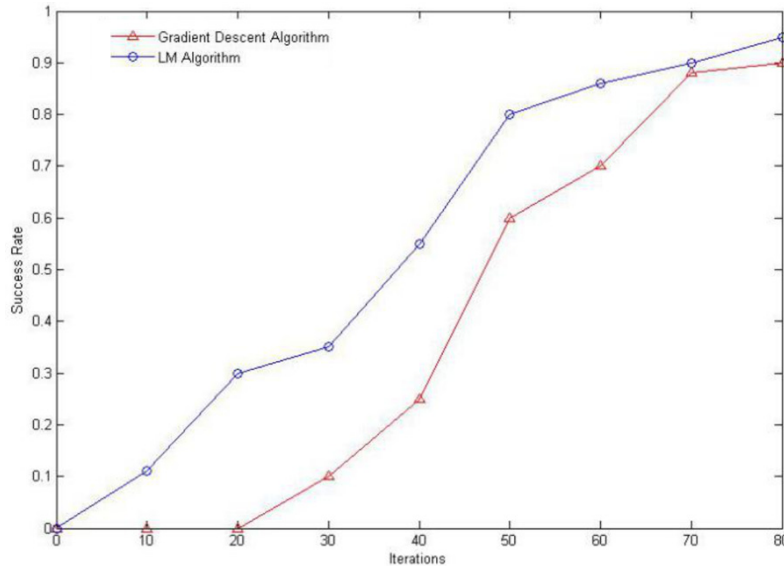


Figure 5. Comparison of iterative steps and prediction accuracy based on two algorithms based on data set D_1 .

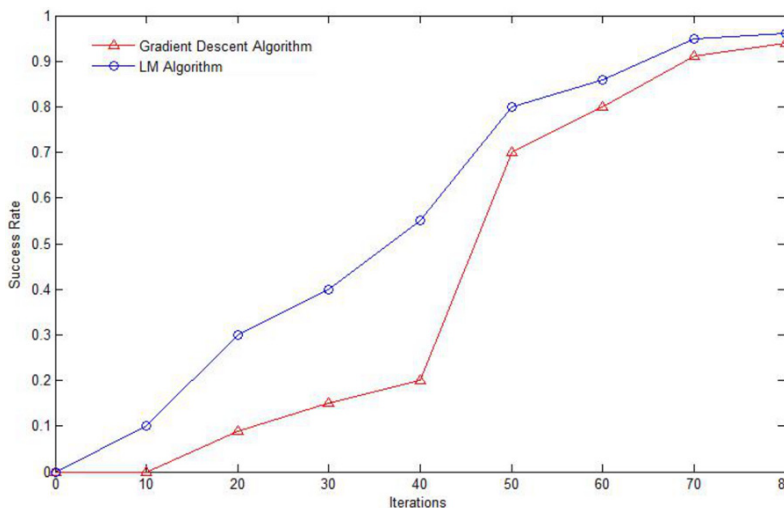


Figure 6. Comparison of iterative steps and prediction accuracy in two algorithms based on data set D_2 .

to the Gauss-Newton method. When the Gauss-Newton method approaches the minimum value of error, the speed of calculation is faster and the accuracy is higher. Since the LM algorithm utilizes the approximate second-order derivative information, it is much faster than the gradient descent method. Experiments have also shown that the LM algorithm can increase the speed by several tens or even hundreds of times compared with the original gradient descent method. In addition, since $J^T(W)$

$J(W) + \mu I$ is positive definite, the solution of equation (6) always exists, and the Gauss-Newton method has requirements for whether or not $J^T(W)J(W)$ is non-singular, so the LM algorithm is also superior to the Gauss-Newton method. In actual operation, μ is a tentative parameter. For a given μ , if the new coefficient W_{k+1} can reduce the error index function E , μ will decrease; otherwise, μ will increase. At the same time, the efficiency of each iteration is significantly improved, which can greatly improve its overall performance, especially when the accuracy requirements are high.

However, the model still has several limitations. First, as can be seen in **Table 1**, the training speed of CNN-LM is slightly inferior to that of CNN constructed using the steepest descent method. This is because the special selection of the long iterative step size of the steepest descent method. Whether or not it can guarantee the other parameters of the method presented in this paper while preserving the recognition accuracy is worthy of further research. Second, the random initialization of the CNN parameters may cause the optimal parameters to fall into a

local minimum value. In addition, the relevant parameters of the convolutional layer and the pooled layer are selected based on human experience. Therefore, it is also a problem worthy of attention in the future to find a suitable method to initialize the network parameters instead of cumbersome manual experiments to determine the optimal parameters. Third, when calculating the Jacobian matrix, the LM algorithm needs to derive partial derivatives for each parameter to be estimated. Therefore,

Table 1. Comparison of different weight update algorithms

	Gauss-Newton	LM	Steepest descent	Gradient descent
Prediction accuracy using D_1 (%)	94.26	95.90	94.26	90.98
Prediction accuracy using D_2 (%)	93.44	96.72	92.62	93.44
Average training time(s)	564	153	149	904

when the objective function is more complicated, the LM algorithm has limitations.

Conclusion

The CNN-LM model proposed in this paper integrates the advantages of the CNN and LM algorithms, and shows better performance in the identification of human diseases of the lower limbs using medical infrared thermograms. Compared with traditional CNN, it has high precision, fast speed, and can be generalized, with a classification accuracy rate of up to 95.90%. Therefore, it can be effectively applied to the diagnosis of human diseases of the lower limbs and allows for the accurate identification of disease conditions ahead of the next treatment, saving time for doctors and patients, effectively avoiding delays in the treatment of the condition. The optimization algorithm proposed in this paper is also applicable to large convolutional neural networks, allowing said network to be more accurate and efficient, with good potential for generalization. It has wide application value in medical image recognition in the future.

Acknowledgements

The authors would like to thank National Key R&D Program of China (2018YFC1311900) and the natural science foundation of Shandong province (ZR2018MF006) for the support to this work. In addition, the authors would like to thank people for agreeing to provide data on infrared thermograms.

Disclosure of conflict of interest

None.

Address correspondence to: Longlong Liu, School of Mathematical Sciences at Ocean University of China, Qingdao 266000, P. R. China. Tel: +86-13173237966; Fax: +86-0532-85901752; E-mail: liulonglongmadi@163.com

References

- [1] 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 Techn* 2017; 85: 315-323.
- [2] Motohiro A, Ueda H, Komatsu H, Yanai N and Mori T; National Chest Hospital Study Group for Lung Cancer. Prognosis of non-surgically treated, clinical stage I lung cancer patients in Japan. *Lung Cancer* 2002; 36: 65-69.
- [3] Krizhevsky A, Sutskever I and Hinton G. ImageNet classification with deep convolutional neural networks. *NIPS* 2012; 25.
- [4] Simonyan K and Zisserman A. Very deep convolutional networks for large-scale image recognition. *Computer Science* 2014.
- [5] Allison MA, Cushman M, Callas PW, Denenberg JO, Jensky NE and Criqui MH. Adipokines are associated with lower extremity venous disease: the San Diego population study. *J Thromb Haemost* 2010; 8: 1912-1918.
- [6] Gregg EW, Gu Q, Williams D, de Rekeneire N, Cheng YJ, Geiss L and Engelgau M. Prevalence of lower extremity diseases associated with normal glucose levels, impaired fasting glucose, and diabetes among U.S. adults aged 40 or older. *Diabetes Res Clin Pract* 2007; 77: 485-488.
- [7] Girshick R, Donahue J, Darrell T and Malik J. Region-based convolutional networks for accurate object detection and segmentation. *IEEE Trans Pattern Anal Mach Intell* 2016; 38: 142-158.
- [8] Zakian C, Pretty I and Ellwood R. Near-infrared hyperspectral imaging of teeth for dental caries detection. *J Biomed Opt* 2009; 14: 064047.
- [9] Lahiri BB, Bagavathiappan S, Jayakumar T and Philip J. Medical applications of infrared thermography: a review. *Infrared Phys Techn* 2012; 55.
- [10] Liu W, Anguelov D, Erhan D, Christian S, Scott R, Cheng-Yang F and Alexander CB. SSD: Single Shot MultiBox Detector. *European Conference on Computer Vision*. Springer, Cham; 2016.
- [11] Schroff F, Kalenichenko D and Philbin J. [IEEE 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) - Boston, MA, USA (2015.6.7-2015.6.12)] 2015 IEEE Confer-

- ence on Computer Vision and Pattern Recognition (CVPR) - FaceNet: a unified embedding for face recognition and clustering. 815-823.
- [12] Ju C, Bibaut A and van der Laan M. The relative performance of ensemble methods with deep convolutional neural networks for image classification. *J Appl Stat* 2018; 45: 2800-2818.
- [13] Liu L, Qu J, Zhou X, Liu X, Zhang Z, Wang X, Liu T and Liu G. Discovery of a strongly-interrelated gene network in corals under constant darkness by correlation analysis after wavelet transform on complex network model. *PLoS One* 2014; 9: e92434.
- [14] Liu LL, Liu MJ and Ma M. Function Clustering Self-Organization Maps (FCSOMs) for mining differentially expressed genes in *Drosophila* and its correlation with the growth medium. *Genet Mol Res* 2015; 14: 11658-11671.
- [15] Liu L, Zhao T, Ma M and Wang Y. A new gene regulatory network model based on BP algorithm for interrogating differentially expressed genes of Sea Urchin. *Springerplus* 2016; 5: 1911.
- [16] Liu L, Ma M and Cui J. A novel model-based on FCM-LM algorithm for prediction of protein folding rate. *J Bioinform Comput Biol* 2017; 15: 1750012.
- [17] Baldeon-Calisto M and Lai-Yuen SK. AdaResU-Net: multiobjective adaptive convolutional neural network for medical image segmentation. *Neurocomputing* 2019.
- [18] Larsson M, Zhang Y and Kahl F. Robust abdominal organ segmentation using regional convolutional neural networks. *Applied Soft Computing* 2018; 70: 41-52.
- [19] Zhao Z, Yang Z, Luo L, Wang L, Zhang Y, Lin H and Wang J. Disease named entity recognition from biomedical literature using a novel convolutional neural network. *BMC Medical Genomics* 2017; 10 Suppl 5: 73.
- [20] Bardou D, Zhang K and Ahmad SM. Lung sounds classification using convolutional neural networks. *Artif Intell Med* 2018; 88: 58-69.
- [21] Aykanat M, Kılıç Ö, Kurt B and Saryal S. Classification of lung sounds using convolutional neural networks. *EURASIP J Image Video Process* 2017: 65.
- [22] Atmane K, Hongbin M and Qing F. Convolutional neural network based on extreme learning machine for maritime ships recognition in infrared images. *Sensors (Basel)* 2018; 18.
- [23] Wang X, Wu YJ, Wang RJ, Wei YY and Gui YM. Gray BP neural network based prediction of rice protein interaction network. *Cluster Computing* 2018.
- [24] Abu-Al-Nadi DI, Ismail TH and Mismar MJ. Interference suppression by element position control of phased arrays using LM algorithm. *AEU-Int J Electron C* 2006; 60: 151-158.
- [25] Nawi NM, Khan A and Rehman MZ. A new Cuckoo Search Based Levenberg-Marquardt (CSLM) algorithm. *Lect Notes Comput Sc* 2013.
- [26] Nguyen-Truong HT and Le HM. An implementation of the Levenberg-Marquardt algorithm for simultaneous-energy-gradient fitting using two-layer feed-forward neural networks. *Chem Phys Lett* 2015; 629: 40-45.
- [27] Zhang K and Wang ZZ. The application of BP neural network improved with LM algorithm in surface EMG signal classification. *Zhongguo Yi Liao Qi Xie Za Zhi* 2005; 29: 399-401.