

# The Detection of Hyperthyroidism by the Modified LeNet-5 Network

Q. ZHANG<sup>1</sup>, J. X. HU<sup>1\*</sup> AND S. ZHOU<sup>1</sup>

Department of Endocrinology, HwaMei Hospital, University of Chinese Academy of Sciences, Ningbo, Zhejiang 315010, China,  
<sup>1</sup>Ningbo Institute of Life and Health Industry, University of Chinese Academy of Sciences, Ningbo, Zhejiang 315010, China

*Zhang et al.: Detection of hyperthyroidism by the modified LeNet-5 network*

To study the automatic detection method of hyperthyroidism based on the deep learning of the modified LeNet-5 network and to establish a detection method with higher precision and better performance, a total of 180 facial images of patients with hyperthyroidism in the ultrasound imaging department of the HwaMei hospital were collected as well as 180 facial images of healthy persons were collected to serve as the control group. A method was proposed for the detection of hyperthyroidism based on the modified LeNet-5. Three test groups were designed by randomized controlled experiment design for the detection of hyperthyroidism, in which the control group was tested manually by experienced doctors, the LeNet-5 network learning group adopted the classical LeNet-5 network learning algorithm and the experimental group used the modified LeNet-5 network learning algorithm. Evaluation indices included the accuracy, detection efficiency, sensitivity, specificity and F1 score of the detection. At the same time, differences between the two algorithms in the detection of thyroid nodules were compared. There was no significant difference between the 3 groups in the accuracy and specificity of detection of hyperthyroidism. In terms of the detection efficiency and sensitivity, the performance of the network learning algorithm group and learning group was better than that of the control group. Both the network learning algorithm group and experimental group could detect the thyroid nodules accurately, but there was no significant difference in the accuracy of detecting the type of thyroid nodules. The correct rate of malignant thyroid nodules was significantly higher than that of the benign thyroid nodules. The modified LeNet-5 network algorithm showed acceptable consistency with experienced doctors in the detection of hyperthyroidism, and this method was useful for the exclusion of thyroid malignant tumours and can be used as a simple method to exclude and identify malignant thyroid tumours.

**Key words:** Hyperthyroidism, LeNet-5 network model, face recognition, automatic detection

The thyroid gland was one of the most important glands in the human body; the thyroid hormones produced by the thyroid gland promote growth and development and regulate metabolism<sup>[1]</sup>. Hyperthyroidism was a clinical syndrome characterized by the excessive secretion of thyroid hormones caused by a variety of causes, which was mainly manifested by increased excitability and hyper metabolism of nervous, circulatory, digestive systems with a typical disease face<sup>[2-5]</sup>. Among them, the Grave's disease (diffuse toxic goiter) was the most common, accounting for about 85 % of all patients with hyperthyroidism<sup>[6]</sup>. Hyperthyroidism was a common endocrine disease; the epidemiological analysis shows that the incidence was 0.14-0.5 %<sup>[7]</sup>. In clinical practice, hyperthyroidism was generally diagnosed by the patient's appearance

and ultrasonic examination<sup>[8-10]</sup>. However, the diagnosis depends on the operator's clinical experience<sup>[6-8]</sup>, which means that less experienced doctors were prone to misdiagnose resulting in an increase in the number of fine needle aspiration (FNA) biopsies. Therefore, it is necessary to develop an automatic detection method of hyperthyroidism to reduce the workload of the doctors, and the dependence on operators, and at the same time to reduce the misdiagnosis rate.

In recent years, the automatic detection methods of hyperthyroidism can be divided into two categories, the artificial feature method and data-driven method. The commonly used artificial feature methods include the local binary pattern (LBP)<sup>[11]</sup>, gray level co-occurrence matrix (GLCM)<sup>[12]</sup> and discrete wavelet transform

\*Address for correspondence

E-mail: m16531752663\_1@163.com

(DWT). However, these methods can only generate low-level information, such as image texture, geometry and statistical distribution and usually need classifiers for classification. Only by giving suitable characteristics for classifying, a proper classification can be arrived at. For data-driven methods, a convolution neural network (CNN) can improve the classification and detection performance of natural images greatly without manual feature description<sup>[13-15]</sup>. In recent years, deep learning has developed rapidly, especially the LeNet-5 network, which was widely used in a variety of computer vision tasks that greatly promoted development of CNN. However, when the LeNet-5 network was used to classify the thyroid and other multi-class targets, the efficiency was not high<sup>[16]</sup>. In order to solve these problems, a new modified LeNet-5 network model was proposed in this investigation based on the analysis of LeNet-5 and other classical CNN. The modified LeNet-5 network introduces the convolution kernel module group, changes the number and size of the convolution kernel appropriately, introduces the batch normalization (BN) layer to standardize the input data, replaces the Sigmoid function with Relu activation function and the full connection layer with global pooling layer. And after experimental verification, the expected outcome was achieved.

## MATERIALS AND METHODS

### The modified LeNet-5 network:

In view of the problems with the traditional LeNet-5 network in identifying thyroid nodules, the following improvements were made. Convolution module groups Inception\_conv3 and Inception\_conv5 replaced the Conv3 and Conv5 of LeNet-5 network, respectively. While extracting more features of the target, increase the depth of the network. The two convolution module groups contain 2-3 convolution layers, respectively. Selection of the convolution kernel was related to whether the effective features can be extracted. The original 5×5 convolution kernel was not effective in the facial feature extraction of hyperthyroidism. Thus the size of the convolution kernel was changed and replaced the original convolution kernel with 2 3×3 convolution kernels in the Inception\_conv3 and Inception\_conv5 modules. The traditional LeNet-5 network has a small number of convolution kernels per layer, which cannot extract the features of the target fully from the face data. Therefore, the number of convolution kernels should be changed reasonably. The introduction of the BN layer can standardize the input batch samples and improve the input of the neural network, which can

improve the training speed of the neural networks to a certain extent and prevent the over-fitting and gradient disappearance. The Sigmoid function adopted by the traditional LeNet-5 was not universal and the Relu function has the characteristic of preventing passing of the negative signal, which was more efficient and can improve the convergence speed. The parameters of the full connection layer account for about 80 % of the network, which greatly increases the network training time. Experiments showed that the full connection layer has no effect on the classification, so the global pooling layer Max\_pool8 was used to replace the full connection layer.

### The detection of thyroid nodule by the modified LeNet-5 network:

The ultrasonic image of the thyroid nodule was the input layer with a size of 64×64. The first layer was Conv1. The size of convolution kernel was 5×5 and the number of convolution kernel was 32 and 32 feature maps with a size of 60×60 can be obtained by Conv1. In the second layer, the pool size of Max\_pool2 was 2×2 and the step size was 2. And 32 feature maps with a size of 30×30 can be obtained by Max\_pool2. The third layer was Inception\_conv3, which has 4 branches, Branch\_0, Branch\_1, Branch\_2 and Branch\_3. Branch\_0 uses 24 1×1 convolution kernels, Branch\_1 uses two cascaded convolution kernels, which were 16 1×1 convolution kernels and 24 5×5 convolution kernels, respectively. Branch\_2 replaces the 5×5 convolution kernels in Branch\_1 with two 3×3 convolution kernels to reduce the number of parameters and extract more abundant features of the target. That is, Branch\_2 consisted of 32 1×1, 64 3×3, and 64 3×3 three-stage convolution kernel cascades and Branch\_3 consisted of a 3×3 mean pool and 16 1×1 convolution cascades and finally, 128 feature maps with a size of 30×30 can be obtained by Inception\_conv3.

The fourth layer was a Max\_pool4, with a maximum pool size of 2×2 and a step size of 2. And 128 feature maps with the size of 15×15 can be obtained by Max\_pool4. The fifth layer was Inception\_conv5, which has the same structure as Inception\_conv3. It also has 4 branches, namely Branch\_0, Branch\_1, Branch\_2 and Branch\_3. The difference was that the number of convolution kernels changed to 64, 64, 96, and 32, respectively and 256 feature maps with the size of 15×15 can be obtained by Inception\_conv5. The sixth layer was Max\_pool6, the pool size is 3×3 and the step size was 3. And 256 feature maps with the size of 5×5 can be obtained by Max\_pool6. In the seventh layer, the

size of Conv7 was 3×3, and the number of convolution kernel was 62. And 62 feature maps with the size of 3×3 can be obtained by Conv7.

The eighth layer was the global pooling layer, and the pooling size was 3×3 and 62 feature maps with the size of 3×3 were pooled to 1×1 and the feature vectors were input to the Softmax layer. The convolution kernel size of Conv8 was 3×3 and the number of convolution kernels was 62. An attempt was made to use the Conv8 in the dotted frame to replace the full connection layer of the traditional LeNet-5 network for network improvement. However, the experimental results showed that the effect of Conv8 was not as good as that of Max\_pool8. The modified LeNet-5 network parameters were shown in Table 1. The surface of the modified network layer was 8 layers, but it was actually 12 layers (excluding input layer and Softmax layer) due to the Inception\_conv3 and Inception\_conv5 were all three layers, which increases the depth of the network.

Facial images of patients with hyperthyroidism were received from the ultrasound imaging department of HwaMei hospital. A total of 180 facial images of patients and 180 images of healthy individuals were collected and a data database of these patients was established.

### Experimental design:

In this investigation, all the image data collected were tested 3 times by experienced clinicians (control group), programs based on LeNet-5 network algorithm (LN group) and programs based on the modified LeNet-5 network algorithm (RLN group), respectively. The 3 tests were carried out randomly and double-blindly (fig. 1). According to the elements of computer recognition, a standard database of disease-related data was established, including the medical history and facial images of patients. The types of diseases including but not limited to the hyperthyroidism, goiter, hypothyroidism, thyroid tumours, and other non-thyroid-related diseases. The system was training and deep learning, including the batch sample training

to the public data resources and the single training to the database data and analyses of the results. On this basis, the empirical parameters set for facial feature recognition of patients with hyperthyroidism were added and the parameters were adjusted to improve the accuracy and sensitivity of the recognition.

### Evaluation indicators and methods:

In the program learning, according to the real categories and the categories predicted by the classification model, the records in the data set could be divided into 4 categories, true positive (TP), false positive (FP), true negative (TN) and false negative (FN). TP means the correct recognition, the positive sample was correctly identified as positive. FP means the wrong recognition, the negative sample was erroneously identified as positive. TN means correct recognition again, the negative sample being correctly identified as negative. FN means wrong recognition again the positive sample being mistakenly identified as a negative. Accuracy was used to measure the proportion of all correct judgments in the classification model, which was defined as (AC),  $AC = (TP + TN) / (TP + TN + FP + FN)$ . Sensitivity was used to measure the correct proportion of model prediction, the true value of all results was positive, it was defined as (PC),  $PC = TP / (TP + FP)$ . Specificity was used to measure the proportion of correct prediction in all results, the true value was negative, it was defined as (ST),  $ST = TN / (TN + FP)$ . F1-Score was a comprehensive index, which was the harmonic average of accuracy and sensitivity. It was defined as (FS),  $FS = 2TP / (2TP + FP + FN)$ . The range of accuracy, precision, sensitivity, and specificity was expressed by a decimal between 0 and 1. The higher the value, the better the classification performance of the model. A recognition control experiment was also conducted on hyperthyroidism patients with thyroid nodules. Two algorithm programs were used for nodule segmentation and nodule diagnosis. Each nodule was segmented once. Nodular segmentation was evaluated visually and was divided into 3 categories. Excellent, the segmented part completely matches the nodule. Satisfactory, but not perfect, the segmented part still represented the nodule. The visual estimate of the maximum contour mismatch between the cover layer and the nodule is no more than 30 % and Poor, some nodules are segmented, but the segmented contours do not represent nodules (estimated mismatch >30 %).

The experiment was based on the Tensorflow deep learning platform under Linux system. The processor was Intel core i7mur7700k, the main frequency was 4.2 GHz, the memory was 16 GB, the video card was

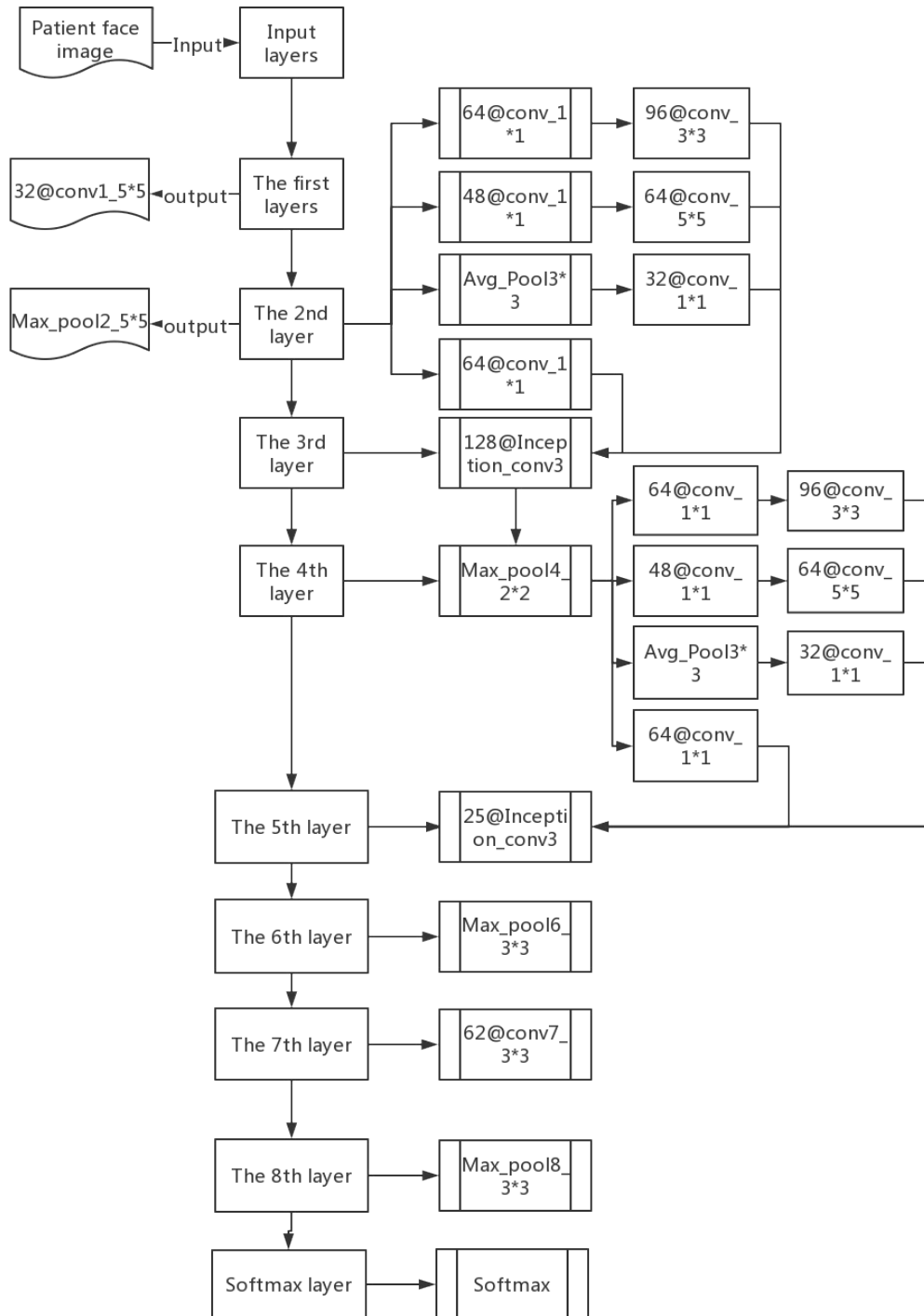
**TABLE 1: TRAINING RESULTS OF DIFFERENT ITERATIONS OF TWO ALGORITHMS**

Iteration times	Training results (%)		p value
	LN group	RLN group	
137	98.334	98.622	0.886
152	98.338	98.513	0.913
201	98.546	98.536	0.996
206	98.664	98.714	0.893
214	98.546	98.742	0.875

NVdia GTX1080TI, and the memory of GPU was 11 GB. In order to give full play to the performance of GPU, the size of Mini-batch was set to 2n (256), the number of Train\_num was set to 1000, and Learn\_rate was set to 0.05. The learning rate was attenuated by the exponential decay method, and the Decay was set to 0.96, weight attenuation of BN was set to 0.005.

### Statistical analysis

The data were processed by SPSS 18.0, the measurement data were expressed as mean±SD. Independent sample t-test was used between groups, paired sample t-test was used within groups, and the counting data were expressed as a percentage and the  $\chi^2$  test was used.  $p < 0.05$  means the difference was statistically significant.



**Fig. 1: The work flow of the thyroid nodule detection by the modified LeNet-5**

## RESULTS AND DISCUSSION

When the LeNet-5 network loads data training, the data set was randomly divided into 80 % training set and 20 % verification set and the network weight was randomly initialized, which will cause small fluctuations in the training results. Therefore, when training the network, different iterative times were trained, and the average value was taken for three times in each iteration. The training results of different iterative times were shown in Table 1. The average training results of the two research groups were about 98.5 %. Although the training results of the two groups were different, there was no statistical difference ( $p > 0.05$ ), indicating that the convergence of sample training tends to be stable and will not fluctuate greatly.

After the training of the 2 algorithm programs, the 3 research groups independently used the data collected for face image recognition. The results showed that although the accuracy of the LN group and the RLN group was lower than that of the control group, there was no statistical difference among the 3 groups (fig. 2 p). The time spent in the LN group and the RLN group was significantly lower than that in the control

group, and the time consumed in the RLN group was the least. It indicated that the screening efficiency of patients with hyperthyroidism can be improved based on the LeNet-5 network algorithm.

As shown in fig. 3, compared to the control group, the sensitivity of the LN group and the RLN group was significantly higher than that of the control group ( $p < 0.05$ ), but there was no significant difference between the LN group and the RLN group. In terms of F1-Score, there was no significant difference among the 3 groups ( $p < 0.05$ ), indicating that LeNet-5 network algorithm could obtain the disease features of patients with hyperthyroidism effectively, especially our modified LeNet-5 network algorithm.

The test results were shown in fig. 4, there was no significant difference between the 3 groups ( $p > 0.05$ ), indicating that the LeNet-5 network algorithm has good accuracy in the detection of hyperthyroidism. As shown in fig. 5, the success rate of the RLN group was 5 % higher than that of the LN group and the success rate of the 2 algorithms for thyroid nodules was 80 %. There was no significant difference between the same types of thyroid nodules ( $p > 0.05$ ). The success rate of

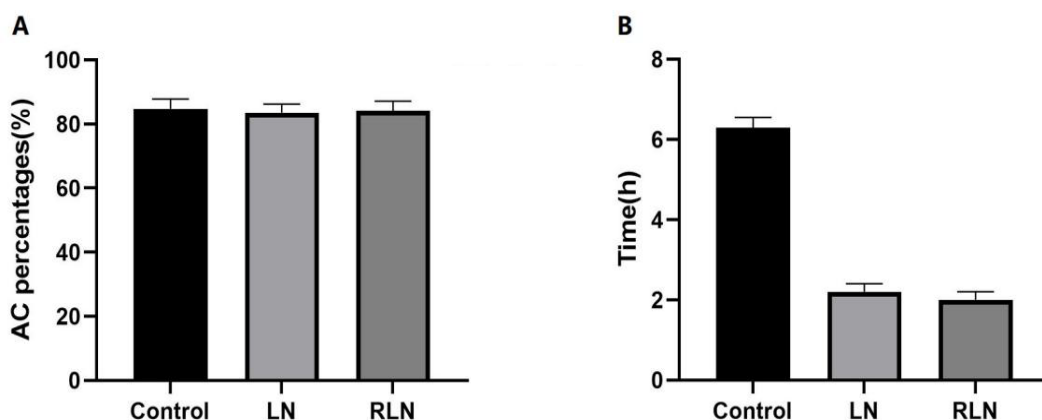


Fig. 2: Accuracy and time required for detection in the three study groups

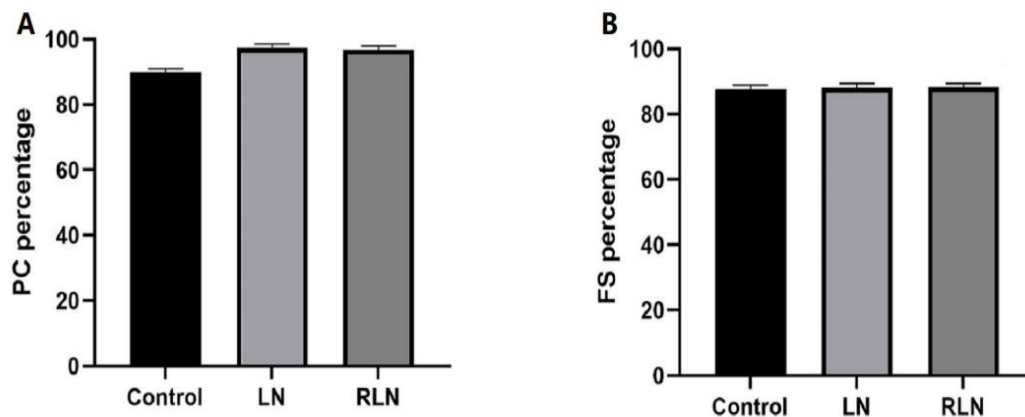


Fig. 3: Differences in sensitivity and F1-Score among the three research groups

the segmentation of malignant thyroid nodules was significantly higher than that of benign thyroid nodules ( $p < 0.05$ ). Excellent, satisfactory and poor segments of benign thyroid nodules were 15.3 (n=9), 66.1 (n=39) and 18.6 % (n=11), respectively. The excellent, satisfactory and poor segments of malignant thyroid nodules were 27.9 (n=12), 67.4 (n=29) and 4.7 % (n=2), respectively ( $p = 0.04$ ).

In this study, the modified LeNet-5 network algorithm for automatic recognition and diagnosis of hyperthyroidism was evaluated. The accuracy and specificity of the modified LeNet-5 network algorithm for the detection of hyperthyroidism were consistent as noted with experienced doctors and the efficiency and sensitivity were higher than that of the experienced doctors. At the same time, the success rate of the modified LeNet-5 network algorithm in the detection of malignant thyroid nodules was significantly better than that of benign thyroid nodules.

In recent years, the face recognition technology has developed rapidly and achieved many research results, which has become one of the important research

directions in the field of computer vision<sup>[17]</sup>. However, because of its complexity, it was easily affected by many factors, such as light, expression, posture. Therefore, the designed algorithm needs to weaken the intra-class changes between the same individuals and strengthen the inter-class changes between different individuals, such as PCA<sup>[18]</sup>, LDA<sup>[19]</sup>, Bayesian Face<sup>[20]</sup> and other face recognition algorithms based on subspace learning. The research of face recognition algorithm based on CNN has made great progress in the past decade. After a lot of data training, these models changed from DeepID, DeepID2, LeNet-5, Face-Net to VGG-Face, these can recognize the features of human face accurately, thus achieving better results in face recognition verification. The LeNet-5 network deep learning method has high accuracy and sensitivity, which accelerates the popularization and application of neural network learning. However, the classical LeNet-5 network algorithm involves many manually modified parameters and complex data algorithm structure, so its efficiency was not high, and the parameter setting depends on experience. Therefore, this investigation made an improvement and introduced the convolution group, which was the core module of the network learning depth algorithm. Firstly, the input facial data of patients were standardized, and then the convolution kernel was modified according to the facial characteristics of patients with hyperthyroidism. Finally the global optimization is carried out by the Relu function to improve the efficiency of solution and recognition in the case of global optimization.

In order to evaluate the efficiency and success rate of LeNet-5 network algorithm in detecting hyperthyroidism, experienced doctors were introduced to diagnose hyperthyroidism. Results obtained showed that there was no significant difference in the detection rate between LeNet-5 network algorithm and the experienced doctors in terms of accuracy or specificity

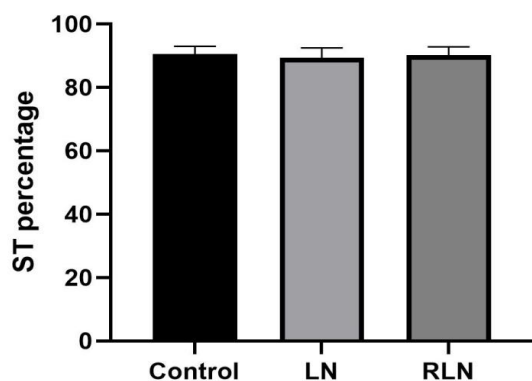


Fig. 4: Differences in the specificity of the three research groups

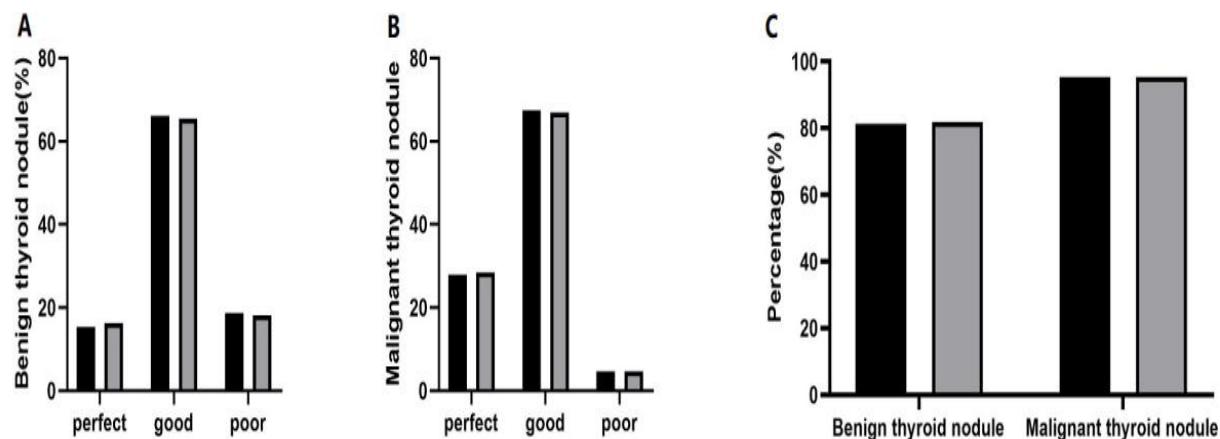


Fig. 5: Detection of thyroid nodules by two algorithms

in the detection of hyperthyroidism, indicating that the LeNet-5 network algorithm was a reliable and fault-tolerant method for the detection of hyperthyroidism. At the same time, the detection efficiency of LeNet-5 network algorithm was about 3 times higher than that of the control group, and its sensitivity to hyperthyroidism was also significantly higher than that of the control group, which showed that LeNet-5 network algorithm could detect hyperthyroidism efficiently and quickly. In order to further verify the superiority and reliability of this modified LeNet-5 network algorithm, thyroid nodules were detected. The detection results showed that the success rate of the improved LeNet-5 network algorithm was higher than that of LeNet-5 network algorithm. At the same time, it was found that the success rate of LeNet-5 network algorithm for the segmentation of malignant thyroid nodules was significantly higher than that of benign thyroid nodules, which indicated that the LeNet-5 network algorithm has the potential to be used to detect malignant thyroid nodules.

In summary, the use of our improved LeNet-5 network algorithm shows acceptable consistency with the experienced doctors in detecting the characterization of hyperthyroidism, and this method was useful for the exclusion of thyroid nausea tumours, and can be used as a simple method to exclude and determine the biopsy requirements of malignant thyroid tumours.

### Acknowledgement:

This study was supported by Zhejiang Medical Science and Technology Research Foundation, China (No. 2018KY697).

### Conflict of interest:

All authors report no conflicts of interest in this work.

### REFERENCES

1. Ma J, Wu F, Jiang T, Zhao Q, Kong D. Ultrasound image-based thyroid nodule automatic segmentation using convolutional neural networks. *Int J Comput Assist RadiolSurg* 2017;12:1895-910.
2. Ma L, Ma C, Liu Y, Wang X. Thyroid Diagnosis from SPECT Images Using Convolutional Neural Network with Optimization. *ComputIntellNeurosci* 2019;2019:6212759.
3. Song W, Li S, Liu J, Qin H, Zhang B, Zhang S, *et al.* Multitask Cascade Convolution Neural Networks for Automatic Thyroid Nodule Detection and Recognition. *IEEE J Biomed Health Inform* 2019;23:1215-24.
4. Lee JH, Ha EJ, Kim JH. Application of deep learning to the diagnosis of cervical lymph node metastasis from thyroid cancer with CT. *Eur Radio* 2019;29:5452-7.
5. Poudel P, Illanes A, Sadeghi M, Friebe M. Patch Based Texture Classification of Thyroid Ultrasound Images using Convolutional Neural Network. *ConfProc IEEE Eng Med BiolSoc* 2019;2019:5828-31.
6. Ko SY, Lee JH, Yoon JH, Na H, Hong E, Han K, *et al.* Deep convolutional neural network for the diagnosis of thyroid nodules on ultrasound. *Head Neck* 2019;41:885-91.
7. Tanaka H, Chiu SW, Watanabe T, Kaoku S, Yamaguchi T. Computer-aided diagnosis system for breast ultrasound images using deep learning. *Phys Med Biol* 2019;64:235013.
8. Tsou P, Wu CJ. Mapping Driver Mutations to Histopathological Subtypes in Papillary Thyroid Carcinoma: Applying a Deep Convolutional Neural Network. *J Clin Med* 2019;8:1675.
9. Zhao Z, Ye C, Hu Y, Li C, Li X. Cascade and Fusion of Multitask Convolutional Neural Networks for Detection of Thyroid Nodules in Contrast-Enhanced CT. *ComputIntellNeurosci* 2019;2019:7401235.
10. Li H, Weng J, Shi Y, Gu W, Mao Y, Wang Y, *et al.* An improved deep learning approach for detection of thyroid papillary cancer in ultrasound images. *Sci Rep* 2018;8:6600.
11. Ma J, Wu F, Jiang T, Zhu J, Kong D. Cascade convolutional neural networks for automatic detection of thyroid nodules in ultrasound images. *Med Phys* 2017;44:1678-91.
12. Halicek M, Little JV, Wang X, Chen AY, Fei B. Optical biopsy of head and neck cancer using hyperspectral imaging and convolutional neural networks. *J Biomed Opt* 2019;24:036007.
13. Zhou H, Jin Y, Dai L, Zhang M, Qiu Y, Tian J, *et al.* Differential Diagnosis of Benign and Malignant Thyroid Nodules Using Deep Learning Radiomics of Thyroid Ultrasound Images. *Eur J Radiol* 2020;127:108992.
14. Xu G, Liu M, Jiang Z, Söfker D, Shen W. Bearing Fault Diagnosis Method Based on Deep Convolutional Neural Network and Random Forest Ensemble Learning. *Sensors Basel*. 2019;19:1088.
15. Bian Y, Wang J, Jun JJ, Xie XQ. Deep Convolutional Generative Adversarial Network (dcGAN) Models for Screening and Design of Small Molecules Targeting Cannabinoid Receptors. *Mol Pharm* 2019;16:4451-60.
16. Zhu H, Fang Q, He H, Hu J, Jiang D, Xu K. Automatic Prediction of Meningioma Grade Image Based on Data Amplification and Improved Convolutional Neural Network. *Comput Math Methods Med* 2019;2019:7289273.
17. Zhang S, Sun F, Wang N, Zhang C, Yu Q, Zhang M, *et al.* Computer-Aided Diagnosis (CAD) of Pulmonary Nodule of Thoracic CT Image Using Transfer Learning. *J Digit Imaging* 2019;32:995-1007.
18. Wang T, Lu C, Shen G, Hong F. Sleep apnea detection from a single-lead ECG signal with automatic feature-extraction through a modified LeNet-5 convolutional neural network. *PeerJ* 2019;7:e7731.
19. Bao R, Yuan X, Chen Z, Ma R. Cross-Entropy Pruning for Compressing Convolutional Neural Networks. *Neural Comput* 2018;30:3128-49.
20. Wei G, Li G, Zhao J, He A. Development of a LeNet-5 Gas Identification CNN Structure for Electronic Noses. *Sensor* 2019;19:217.

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms

This article was originally published in a special issue, "Biomedical Research in Healthcare Setting" *Indian J Pharm Sci* 2020;82(3)Spl issue5;110-116