Journal of Computers Vol. 34 No. 5, October 2023, pp. 205-212
doi: 10.53106/199115992023103405016

Intelligent Diagnosis Method for Orthopedic Diseases Based on Medical Images

Wei-Fang Li*, Hai-Xin You, Xiao-Yu Guo

North China University of Science and Technology,
Tangshan City 063210, Hebei Province, China
{lwf897243, hx987767, gxy98789}@126.com

Received 1 August 2023; Revised 1 September 2023; Accepted 28 September 2023

Abstract. This article proposes an orthopedic image recognition method based on an improved convolutional neural network to address the traditional diagnostic methods of orthopedic diseases and the diverse types of orthopedic diseases in the diagnostic process. This method uses a fixed convolutional number to extract key features of orthopedic diseases and reduce the number of features. The article takes the diagnosis of gout in bone diseases as an example and designs an evaluation method based on visualization technology and quantitative indicator calculation. The quantitative indicator calculation obtains the total volume information of urate crystals, thereby assisting doctors in gout diagnosis. The experimental results show that the diagnostic accuracy reaches 98.7%, which can meet the actual diagnostic requirements.

Keywords: convolutional neural network, Intelligent diagnosis, gout

1 Introduction

There are various types of orthopedic diseases and complex professional backgrounds. Usually, doctors use the naked eye to diagnose orthopedic medical images based on their own experience during the diagnosis process. With the innovative development of computer, vision, and artificial intelligence technology, computer-aided systems and intelligent robot technology have been used in various orthopedic surgeries, including arthroplasty, spinal surgery, bone tumor surgery, arthroscopy, fracture reduction Fracture fixation during trauma surgery, etc. However, in the diagnosis of orthopedic diseases, computer-aided technology needs to be strengthened.

At present, the diagnosis of orthopedic diseases mainly relies on medical images. Traditionally, doctors use two-dimensional medical images to identify abnormal locations of lesions and imagine reconstructing their size and shape in the brain. This diagnostic method lacks quantitative indicators. This article focuses on common orthopedic diseases such as gout and cartilage tissue defects, The research on medical image recognition methods involved in computer-aided gout diagnosis and cartilage defect diagnosis is as follows:

1) Firstly, convolutional neural networks are used to diagnose orthopedic diseases. The original network is improved by replacing the backbone network, adding attention modules, and increasing the number of convolutional layers, making the improved network more suitable for the diagnosis of orthopedic diseases.

2) Quantifying the diagnosis of gouty bone disease

3) Build a simulation platform, use a database to train the network, and diagnose and test gout diseases.

In order to describe the above process clearly, the structure of this article is as follows: Chapter 2 mainly introduces some research achievements of relevant scholars and conducts some comparative analysis. Chapter 3 mainly improves the existing convolutional neural network and describes the improvement methods in detail. Chapter 4 quantitatively evaluates the gout diagnosis indicators based on the current situation of gout diagnosis. Chapter 5 is the simulation experiment and result analysis section, and Chapter 6 is the conclusion section, Summarized the innovative points of the article and elaborated on further research.

2 Related Work

Scholars at home and abroad have conducted extensive research on the application of artificial intelligence in

* Corresponding Author

disease diagnosis. Seung uses convolutional neural networks for facial bone diagnosis to determine whether orthognathic surgery is necessary based on facial photos. Through recognition of 822 patient photos, CNN is able to determine whether surgery is necessary solely through photos, with an accuracy rate of 91.2% [1]. Based on X-ray images, Shintaro uses deep convolutional neural networks combined with transfer learning strategies to classify and label various dental implant systems. In practical applications, VGG16 has the highest accuracy in dental implant system classification and is worth promoting [2]. Dong Han studied the latest research and application progress of artificial intelligence in the field of medical imaging, and analyzed the prospects of artificial intelligence technology in medical applications, pointing out the direction for intelligent diagnosis [3]. Tingting Li, based on the entire wrist X-ray digital radiography (DR) image, used deep learning for bone age assessment. After diagnostic experiments on 858 real images aged 0-18 years, the results showed that the AI-BAA model with higher-order features of wrist DR images is a reliable intelligent diagnostic method [4]. The artificial intelligence of the tea ceremony, the artificial intelligence image, and the medical treatment of the high school medicine. The VGg model is an image of the X-rays of the lung [5].

3 Establishment of Improved Neural Network Model

This article uses the updated YOLOv4 as the basic algorithm. In order to make the algorithm structure more lightweight, the Ghostnet network is used instead of the original backbone network, and deep separable convolutions are used instead of the neck network. At the same time, the convolution layers before and after spatial pyramid pooling are increased to 5 layers to increase recognition speed [6].

3.1 Replace Backbone Network

Ghostnet is a convolutional network designed by Huawei, using Ghostnet as the backbone network of YOLOv4. By using Ghostnet, the original network's feature extraction capability can be achieved with fewer parameters and computational complexity [7]. The Ghostnet network contains a Ghost module inside, which first reduces the number of channels and then performs convolutional feature extraction. This running network structure greatly reduces the computational load and increases the computational speed. The structural diagram of the Ghost module is shown in Fig. 1.

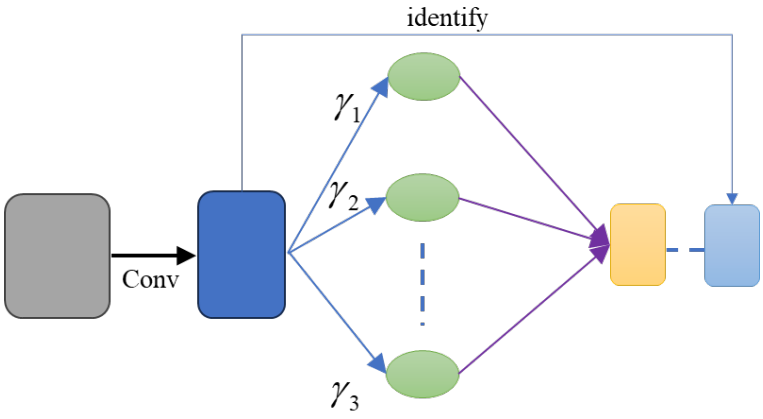

Fig. 1. Schematic diagram of Ghost module structure

The input of the feature map is $H_{in} \times W_{in} \times C_{in}$ (representing the height, width, and number of channels of the feature map), the convolution kernel is $K \times K$ (representing the height and width of the convolution kernel), and the output is $H_{out} \times W_{out} \times C_{out}$. Among them, d (representing the height and width of the convolutional kernel) is equal to K, indicating that it has undergone s transformations. The parameter quantity of Ghost convolution is expressed as:

$$params = \frac{n}{s} \times C_{in} \times K \times K + (s-1) \times \frac{n}{s} \times d \times d. \tag{1}$$

The computational complexity of Ghost convolution is expressed as:

$$FLOPs = \frac{C_{out}}{s} \times H_{out} \times W_{out} \times C_{in} \times K \times K + (s-1) \times \frac{C_{out}}{s} \times H_{out} \times W_{out} \times d \times d. \tag{2}$$

From the above formula, it can be seen that the computational complexity of Ghost convolution has been reduced.

3.2 Add Attention Module

The SENet attention module is used in the Ghostnet network. The SENet attention module is designed to obtain a global receptive field, i.e. a larger field of view. SENet is divided into two steps: compression and excitation. Compression is the process of compressing dimensions through pooling, which can reduce overall computational complexity and better extract features [8].

3.3 Replacement Depth Separable Convolution

In order to further simplify the network, improve its computational speed, and shorten the prediction time of the network for the target, deep separable convolutions are used to replace the original convolution structure. Deep separable convolution consists of channel wise convolution and point wise convolution [9]. Each channel wise convolution has three 3×3 convolution kernels, each corresponding to a channel, and each channel corresponding to a single convolution kernel is convolved to obtain the corresponding Maps. Three 3×3's correspond to three Maps. Channel by channel convolution utilizes each convolution kernel corresponding to a channel, greatly reducing the amount of parameters and computation, and can reduce the complexity of the network. Deep separable convolution is shown in Fig. 2.

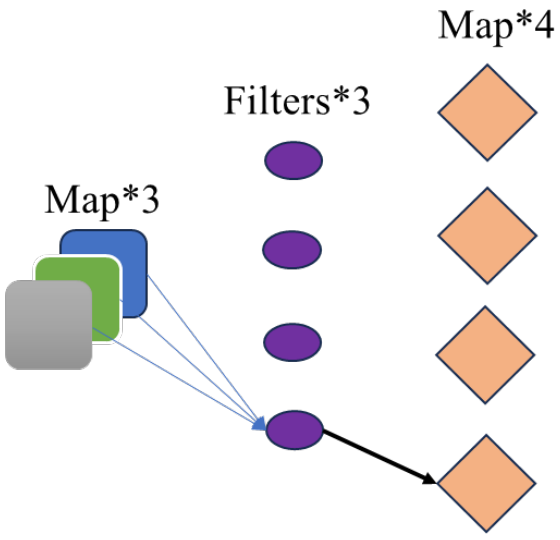

Fig. 2. Schematic diagram of point by point convolution

Channel by channel convolution is usually followed by point by point convolution. Point by point convolution is very similar to ordinary convolution, except that the ordinary convolution kernel is replaced by 1×1×*M*'s convolution kernel, where *M* is the number of channels in the previous channel. Point by point convolution uses 1×1×*M*'s convolution to check the original image for point convolution, thus allowing for channel information fusion through point by point convolution of the image. Channel by channel convolution is the convolution of corresponding channels, without fusing the information of each channel, which can cause information asymmetry. By using point by point convolution, the information of each channel can be fused, which can balance the information of each channel.

When the input of the feature map is $H_{in} \times W_{in} \times C_{in}$, the convolution kernel is $K \times K$, and the output is $H_{out} \times W_{out} \times C_{out}$, the number of convolution parameters per channel in depth-separatable convolution is:

$$params = K \times K \times C_{in}. \tag{3}$$

The computational complexity of channel by channel convolution in deep separable convolution is:

$$FLOPS = K \times K \times C_{in} \times H_{out} \times W_{out}. \tag{4}$$

Number of pointwise convolution parameters in deep separable convolution:

$$params = C_{in} \times C_{out}. \tag{5}$$

The computational complexity of point by point convolution in deep separable convolutions:

$$FLOSPS = C_{in} \times H_{out} \times W_{out} \times C_{out}. \tag{6}$$

3.4 Add Convolution Layers

There is a set of *CBL**5 at the front and back of the spatial pyramid pooling module to further extract image feature information, which can strengthen the feature extraction and fusion of image core information. The purpose of adding convolutional layers is to increase the accuracy of the entire network. After the number of convolutional layers is increased, a new network structure is generated. Therefore, the modified YOLOv4 overall network process is shown in Fig. 3.

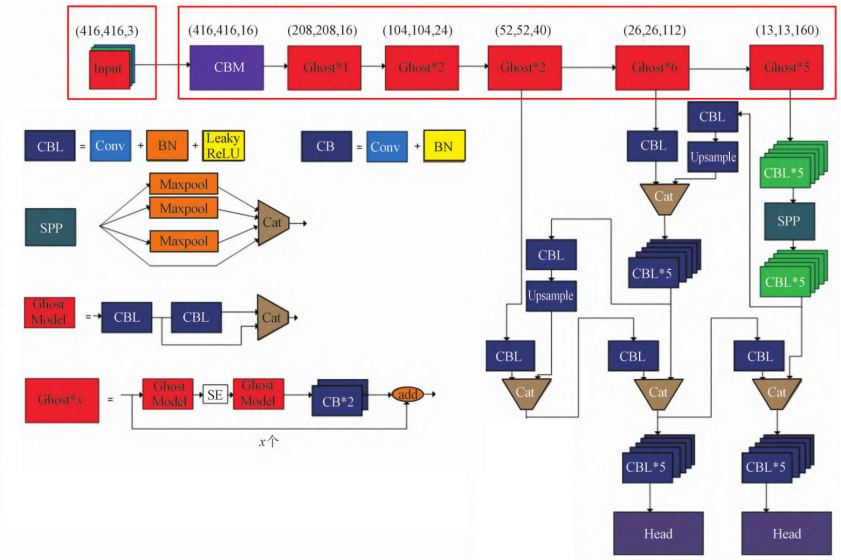

Fig. 3. Improved network structure

4 Quantitative Index Evaluation Method for Gout

Medical image visualization can more intuitively present useful information related to bone disease diagnosis. Gout is located in the urate crystal area. In order to achieve better accuracy in computer-aided diagnosis, this article quantifies the evaluation of urate crystal areas.

In the diagnosis of gout, the total volume of urate crystals is an important indicator reflecting the state of gout. Whether the patient has a gout attack and the clinical staging of gout are positively correlated with the total volume of urate crystals. In addition, the changes in the total volume of urate crystals during different treatment stages can also be used to track and evaluate the efficacy achieved by patients and their post treatment status. The area with the highest deposition of urate crystals is located at joints such as the ankle, inter tarsal, and metatarsophalangeal joints, which cannot be reflected from the total volume of urate crystals. Therefore, quantification of indicators is particularly important. The pseudo code representation of the quantitative evaluation method for urate is as follows:

Algorithm 1. Calculation method for quantitative indicators of total volume of urate crystals

1: Input: Image $I = x, y, z, L_i\ i \in [0, N]$, processed by bone and urate crystal region segmentation algorithm and urate crystal artifact removal algorithm, with physical pixel spacing X_p, Y_p, and Z_p in three coordinate directions. Among them, x, y, and z represent coordinates, L_i represents the class marker values of the pixel points ($L_i = 0$ represents urate crystals, $L_i = 1$ represents bone).
2: Output: Total volume V_M of urate crystals.
3: Process:
4: Initialize the total number of pixels $N_u = 0$ for urate crystallization;
5: Calculate the volume $V_p = X_p * Y_p * Z_p$ of a single pixel;
6: Whiel Pixel point B in image A has not been traversed completely **do**
7: if $P.L_i = 0$ **then**
8: $N_u + +$ **;**
9: end if
10: end while
11: Calculate the total volume V of urate crystals.
12: End

5 Using an Improved Neural Network for Gout Assisted Diagnosis Experiment

This article takes PC as the platform, and the configuration of the PC is Intel (R) Core (TM) i7-9750H CPU @ 2 60GHz 2.59GHz, GPU is NVIDIA GeForce GTX 1660 Ti. The system adopts Win10, the network programming environment is Python 3.6, and Python version 1.7.2 is used.

5.1 Dataset and Model Training

The experimental data was collected from a certain orthopedic hospital in Tangshan City. After flipping correction and color adjustment, the overall dataset consisted of 896 photos. When training the neural network, the input image pixel is 416×416, and the photos are allocated to the training set, validation set, and test set in a 6:1:1 ratio. The clustering algorithm is used to calculate a prior box suitable for this network model. In order to verify the performance of the proposed network, this article uses three indicators: recall, accuracy, and average accuracy to evaluate the network performance (*MAP*). The expression for recall rate is:

$$recall = \frac{TP}{TP + FN}. \tag{7}$$

The expression for accuracy is:

$$precision = \frac{TP}{TP + FP}. \tag{8}$$

In the formula, *TP* is the amount of positive samples detected in positive samples, *FP* is the amount of positive samples detected in negative samples, *FN* is the amount of negative samples detected in negative samples, *MAP* is an important indicator for measuring a network, and the formula for calculating the value of *MAP* is:

$$MAP = \sum_{i=0}^{n} AP_i. \tag{9}$$

To demonstrate the reliability of the improved algorithm proposed in this article, YOLOv4, YOLOv5, SSD, and improved YOLOv4 were used respectively. For convenience in description, improved YOLOv4 is referred to as i-YOLOv4. Comparative experiments were conducted in the same experimental environment configuration and dataset. When inputting pixels with image size A, the speed, corresponding accurate values, and corresponding weights of each network were compared and analyzed. The comparison results are shown in Fig. 4.

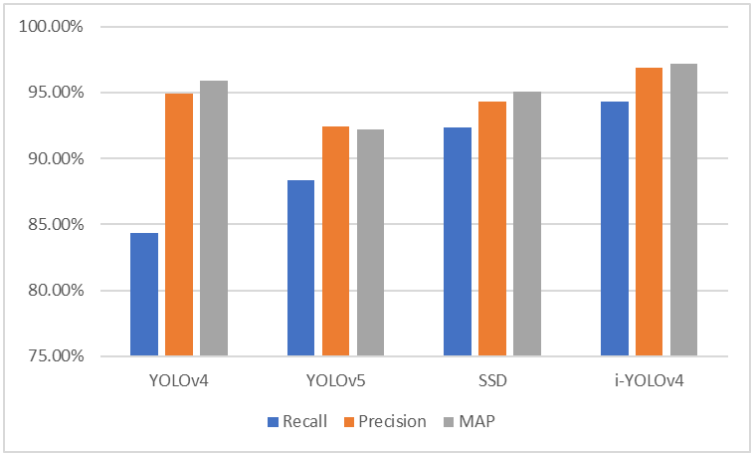

Fig. 4. Model comparison results

The modified network, compared with the first three networks, shows a small difference in MAP values between the modified network and the original YOLOv4 algorithm model, with a difference of 1.3%. Although the accuracy improvement in practical applications is not high, within an acceptable range, the accuracy is higher compared to the other two models. In terms of recall rate, the modified network algorithm model is significantly better than other models. The recall rate reflects the speed of the algorithm, so the improved algorithm can achieve real-time detection. Compared to the YOLOv4 algorithm model, it greatly accelerates the detection speed. Overall, the overall performance of the modified network is better than the original YOLOV4 algorithm model, and is more accurate and fast than the other two networks, which can be applied in practice.

5.2 Analysis of the Effect of Gout Diagnosis Experiment

Using the improved YOLOv4 model on a computer for gout diagnosis experiments, in order to verify the accuracy and effectiveness of the gout auxiliary diagnosis function of this system, this section uses actual cases to assist doctors in gout diagnosis on the orthopedic disease computer-aided diagnosis system in this article. The gout patient case comes from the Medical Imaging Center of a certain orthopedic hospital in Tangshan. The left dorsum of the foot is swollen and accompanied by pain. Dual energy CT detection was performed on a Siemens CT machine to obtain CT sequence images at 80KV and 140KV. The specific information is Row = 496, Column = 496, Pieces number = 315, Pixel Spacing = 0.486mm, and slice spacing=1.0mm. The image example is shown in Fig. 5.

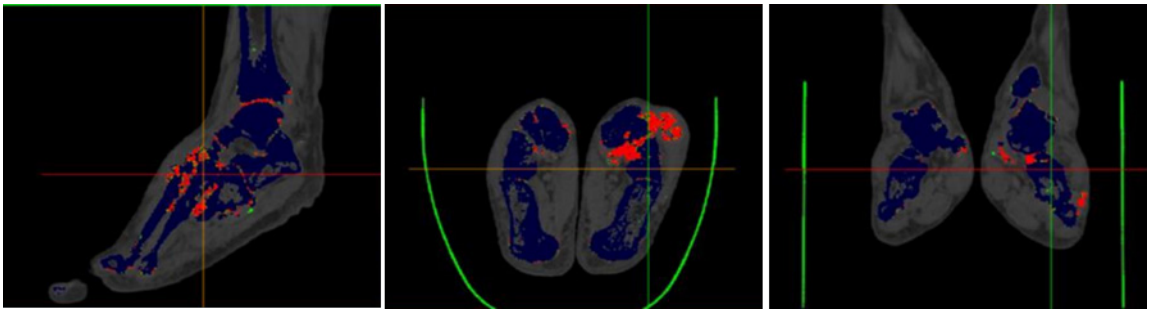

Fig. 5. Gout image

The diagnostic results are shown in Fig. 6, which indicate the labeling and probability of gout symptoms.

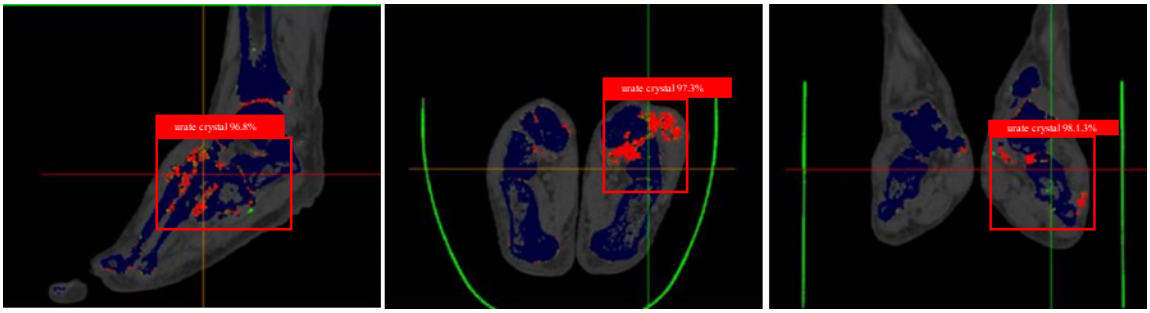

Fig. 6. Recognition result

6 Conclusion

This article uses convolutional neural networks to diagnose common gout in orthopedic diseases, mainly using an improved model to self identify urate crystals in X-ray two-dimensional images. In order to achieve high accuracy in bone disease diagnosis, this article improves the traditional YOLOv4 network model by adding attention mechanisms and increasing the number of convolutions to meet the needs of bone disease diagnosis. Finally, the effectiveness of the proposed method is confirmed through experiments.

At the same time, this article also has the following shortcomings:

1) The database used for training the model is not yet rich enough, and more relevant data needs to be collected in the future

2) This article currently focuses on the diagnosis of gout. There are many types and complexities of bone diseases, and further in-depth research will be conducted in the future to strive for the diagnosis of more bone diseases.

References

[1] S.-H. Jeong, J.-P. Yun, H.-G. Yeom, H.-J. Lim, J. Lee, B.-C. Kim, Deep learning based discrimination of soft tissue profiles requiring orthognathic surgery by facial photographs, Scientific reports 10(1)(2020) 16235.

[2] S. Sukegawa, K. Yoshii, T. Hara, K. Yamashita, K. Nakano, N. Yamamoto, H. Nagatsuka, Y. Furuki, Deep Neural Networks for Dental Implant System Classification, Biomolecules 10(7)(2020) 984.

[3] D. Han, Q.-H. Li, W. Cai, Y.-W. Xia, J. Ning, F. Huang, Research and application of artificial intelligence in medical imaging, Big Data Research 5(1)(2019) 39-67.

[4] T.-T. Li, X.-J. Yang, Q. Wang, X.-H. Ren, J. Lan, G.-J. Yu, B. Li, L.-H. Li, Y. Wen, X. Chen, Study on Artificial

Intelligence Bone Age Assessment Based on Deep Feature Learning for the Whole Hand Digital Radiograph, China Digital Medicine 14(11)(2019) 29-33.
[5] D. Xie, F. Zheng, T. Chen, W.-J. Liu, Remote Diagnosis Algorithm and Implementation of Medical Image Based on CNN, Computer Knowledge and Technology 17(23)(2021) 137-139.
[6] Q.-Q. Ge, Z.-J. Zhang, L. Yuan, X.-M. Li, J.-M. Sun, Safety helmet wearing detection method of fusing environmental features and improved YOLOv4, Journal of Image and Graphics 26(12)(2021) 2904-2917.
[7] C.-J. Liu, X.-P. Duan, B.-W. Xie, Improvement of ECO target tracking algorithm based on GhostNet convolution feature, Laser Technology 46(2)(2022) 239-247.
[8] C. Lin, Z.-S. Yin, Attention-based algorithm for image blind deblurring, Journal of Computer Applications 40(S2) (2020) 151-157.
[9] L.-H. Lin, Y. Wei, J.-H. Pan, Classification of Wuyi rock tealeaves based on convolutional neural network, Journal of Ningde Normal University (Natural Science) 33(4)(2021) 363-369.