A Comparative Study of CNN and DBN Algorithms in Oral Disease Imaging Diagnosis

Kexin Wang, Guangyan Wang*, Aihemaiti Gulibusitan, Ziming Wei, Yuxin Xu, Jinyuan Zhang
School of Information Engineering, Tianjin University of Commerce, Tianjin, China
**Corresponding Author.*

Abstract: Oral imaging involves imaging techniques and diagnostic methods of the internal structure of the oral cavity. Intelligent diagnostic techniques based on machine learning and medical gold standards have become a research hotspot in recent years. In order to obtain the accuracy of deep learning for oral image segmentation, this paper mainly compares the performance of deep belief network (DBN) and convolutional neural network (CNN) in image segmentation and recognition. Firstly, the CBCT image data set of the teeth of the relevant cases was established, and the image was preprocessed and marked, and the training set, test set and verification set were divided. Secondly, the DBN and CNN deep learning neural network models are built and debugged for training and testing respectively. Finally, the performance of the two models is compared and their advantages and disadvantages are analyzed from the data of graphic distortion, feature extraction accuracy, tooth discrimination and accuracy. The experimental results show that both of them have good performance in image segmentation, but the CNN model is superior to DBN in oral image classification and recognition. This study explores the application of deep learning in CBCT image processing of teeth. The research results are of positive significance for assisting doctors to judge the condition more accurately, improving the treatment effect and quality, and improving the digitization and intelligence of oral disease imaging diagnosis.

Keywords: Medical Image Processing; Oral Diseases; Deep Learning; CNN; DBN

1 Introduction

The research of Computer Aided Diagnosis (CAD) [1] shows that machine learning can diagnose the abnormal parts of oral cavity in image diagnosis, which reduces the error rate of traditional machines and helps doctors analyze the diagnosis images and the causes. Oral medical image information is the most direct information source in the diagnosis of oral diseases. Doctors can quickly diagnose the condition according to the content of image diagnosis and carry out the next treatment for patients. In addition, oral imaging can also be used for preoperative and postoperative evaluation and follow-up, so as to better use oral imaging for diagnosis and treatment.

At present, one of the research hotspots in the field of medical imaging is to use deep learning to segment, identify and process images [2], in which computer-aided doctors play a very important role in image diagnosis. There are more or less problems in traditional medical testing instruments, such as low spatial resolution of images, easy to miss diagnosis, inaccurate examination results due to radiation images [3], and there are obvious differences in doctors' testing techniques. In traditional medical imaging diagnosis, there is often the risk of missing diagnosis or wrong diagnosis. Therefore, the popular machine vision technology is applied to the diagnosis of oral diseases. In recent years, the medical field has introduced deep learning models to assist doctors in diagnosis [4] diseases is even higher than that of professional dentists, which greatly reduces the workload of doctors.

Image processing is one of the core problems in the field of computer vision. Convolution neural networks (CNN) and deep belief network (DBN), as the mainstream technologies of image processing, show strong robustness and universality. Zhao et al.[5]

proposed a hybrid structure based on CNN and local image features to achieve FPV pedestrian navigation. Kaur et al.[6] proposed an image fusion framework based on DBN, which further improved the performance of image fusion process. A large number of experiments show that the proposed technology is superior to the existing image fusion technology. Tang et al.[7] proposed a new framework that combines CNN and DBN to classify charts. Compared with the previous methods using original feature extraction, the deep feature has better scalability and stability, and it is proved that the proposed framework is much better than the existing methods. Yang et al.[8] introduced a method based on CNN-LSTM as the backbone of computer vision-based vibration measurement technology. The robustness of the deep learning model is tested by using samples of various materials. Different sizes and results show a high level of perceptual accuracy. At present, DBN and CNN are widely used in the imaging diagnosis of oral diseases. Among them, DBN is an unsupervised learning method, and the loss function in the process is reconstruction error. CNN is used in supervised learning. It can learn the weight of the network through supervised signals, and at the same time, it can get vector coding, which can extract complex oral image features more effectively.

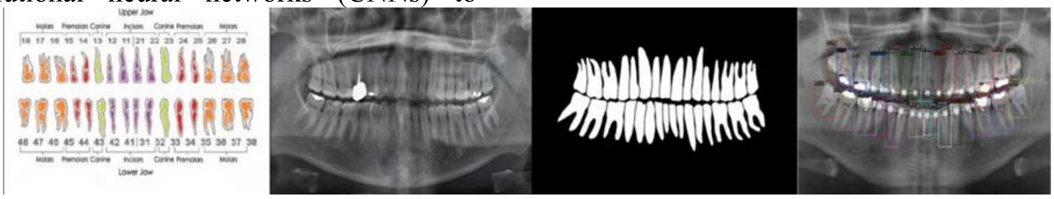
CNN has made a lot of progress in the application of medical image recognition. Tuzoff etal.[9] studied a system based on CNN to automatically detect tooth images and automatically number them. More than one thousand tooth slices were used to simulate and train the system. Deep learning was used to label the surroundings of the teeth in a given tooth slice and classify the teeth. Heuristic algorithm was used to improve the results. The experimental results were compared with the diagnostic results of the physician. It was found that the accuracy and resolution of the system were similar to the diagnostic conclusions given by the experts. Poedjiastoeti et al. [10] developed a CNN model for the detection and identification of ameloblastoma and odontogenic keratocyst in panoramic radiographs. The experimental results show that the results of CNN diagnosis are very similar to those given by dentists, but the time used by dentists is far more than that of machine diagnosis, which reflects the high availability, sensitivity and accuracy of CNN. Choiet al.[11] used a set-up CNN system to examine the probability of dental caries generated by a given tooth piece, and the results were similar to the predictions. Orhan [12] developed a CNN system to diagnose periapical periodontitis in patients. The results showed that the accuracy of artificial intelligence system detection was as high as 92.8 %, which also showed that the CNN learning model could be used as a detection tool for periapical lesions with high accuracy.

In recent years, DBN has been widely used in the field of medical imaging diagnosis. In 2011, Mohamed et al. [13]proposed continuous discriminant training criteria to optimize the weights and state change parameters of deep belief networks. They used Mel frequency cepstrum coefficients and Mel scale filters to train deep belief networks to generate high-level restricted Boltzmann machine features. After updating with the back-propagation method, the image performance recognized on the data set is significantly better than other methods. In the same year, Dahl [14] proposed a context-dependent ' deep belief network-hidden Markov model (DBN-HMM), which was tested on a large number of image recognition datasets. The results obtained are significantly better than the ' Gaussian mixture-hidden Markov model (GMM-HMM)'. It can be seen that the research of CNN and DBN in the recognition and diagnosis of dental medical images has achieved good improvement compared with traditional methods. Based on the self-built dental CBCT data set, this paper will construct CNN and DBN models respectively, compare the experimental results obtained by CNN and DBN, and evaluate the performance difference of two deep network learning in the image diagnosis of oral diseases.

2 CNN and DBN Model Structure and Algorithm

2.1 Principle of CNN Algorithm

CNN is a commonly used deep neural network model or a multi-layer perceptron similar to artificial neural networks, which is often used to analyze visual images [15]. Creating a suitable neural network can make it have a simple judgment function like a human, so that image recognition can have a better effect. As can be seen from Figure 1[16], the input tooth image is sent into the convolutional neural network and extracted through the convolution layer. After that, the maximum pooling filter details are generally used, and finally the feature is expanded in the full connection layer. Lee et al [17] used a variety of migration learning methods and combined with deep convolutional neural networks (CNNs) to monitor osteoporosis in dental panoramic x-rays (DPRs). Some migration learning strategies will affect the deep CNN model, including basic CNN3 and visual geometry group 16 (VGG-16). VGG-16 is better because migration learning and fine-tuning improve the overall effectiveness of deep CNN in screening for osteoporosis in DPRs.



(a)Teeth ISO Numbering (b)Original Teeth Image (c)Teeth Segmentation Mask (d)Teeth Identification Annotation

Figure 1. Dental Recognition based on Convolutional Neural Network

CNN usually includes: input layer, convolution layer, pooling layer and fully connected layer, each with different functions. The structure is shown in Figure 2. These layers can be superimposed to form a complete CNN structure.

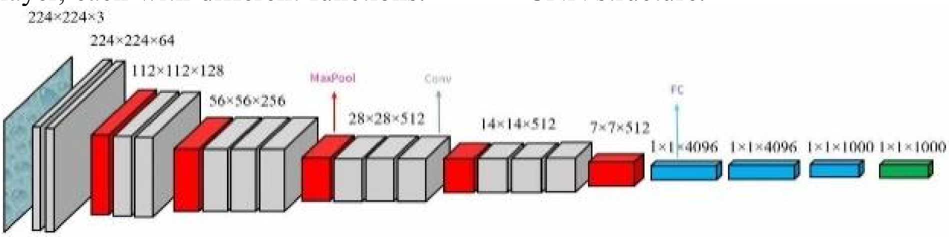


Figure 2. Convolutional Neural Network Diagram

Convolutional Layer: Convolutional Layer is the core layer of CNN. The main function is to extract information from the input picture, which is the image characteristics, which are represented by the pixels in the image in combination. The processing of the convolution layer is the convolution operation. Through the convolution check, the matrix of each channel is carried out from left to right (the convolution kernel is generally a 3 × 3 matrix) from top to bottom (first from left to right, then from top to bottom, so the convolution operation will also retain the position information). The convolution formula is

$$g(x,y) = \omega * f(x,y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} \omega(s,t) f(x-s, y-t) \quad (1)$$

Among them, g(x,y) is the processed matrix, w is the convolution kernel, $f(x,y)$ is the original matrix, $-a \le s \le a$ and $-b \le t \le b$. Slide step by step from the upper left corner to the lower right corner. The sliding step length is a super parameter. Add the values of the three channels to get the output of the convolution operation. The calculation formula is

$$y = f\left(\sum_{i=1}^{n} w_i x_i + b\right) \quad (2)$$

Among them is the weight of the neuron, which is bias.

Pooling layer: Max Pooling Layer's main function is downsampling. The pooling operation can fuse the feature map of the previous layer, retaining the main features while reducing parameters and calculations, because the parameters of adjacent areas have strong relevance and can also prevent the occurrence of fitting. After the feature extraction of the convolution layer, the output feature graph will be transmitted to the pooled layer for feature selection and information filtering. The calculation formula is

$$y_j^l = f\left(\beta_j^l \,\mathrm{down}(x_j^{i-1}) b_j^l\right) \quad (3)$$

Among them, and represents the corresponding weight and bias respectively, and down(·) represents the downsampling function.

Fully Connected Layer: The main function is to realize classification. Each layer in the full-connected layer is a tile structure composed of many neurons. In the full-connected layer, the output of the last convolution layer is flattened and each node of the fully connected layer is connected to each node in the previous layer. Combine the output characteristics of the previous layer, so the weight parameters of this layer are the most. The function of the fully connected network is to stretch the feature map obtained by the last layer of convolution into a vector, multiply the vector, finally reduce its dimension, and then enter it into the softmax layer to get the corresponding score of each category.

When learning neural networks, due to the large amount of data and other reasons, when using CNN for image processing, gradient operations are generally performed many times to achieve the purpose of efficient processing medical images. Due to the diversity of pathology, when using CNN model diagnostic images, it is necessary to have a deeper understanding of specific scenes, such as pathological medical principles, characteristic results, data analysis, etc. At present, the CNN model cannot achieve automatic adaptation and requires manual training or pretreatment, and CNN cannot take into account the accuracy and efficiency of diagnostic images. Despite the above problems, the CNN model still obtains better results in processing medical image data than traditional models.

2.2 Principle of DBN Algorithm

DBN is a probabilistic generative model, which is composed of multiple unsupervised learning Re-structed Boltzmann Machine (RBM) layers and the last layer of SoftMax classifier. Its structure is shown in Figure 3. DBN has multiple hidden layers, which can perform higher-level correlation operations on data. This also makes the hidden layer before and after training have more order. After the previous hidden layer is fully trained, the next hidden layer can perform the same type of training again through the training data obtained by the previous hidden layer, and continuously repeat this process to finally obtain the trained data.

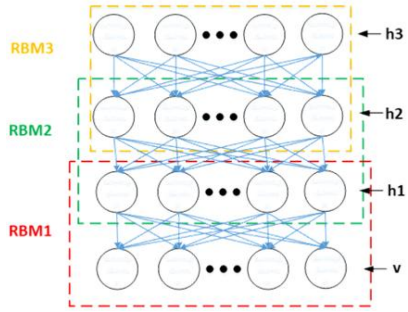


Figure 3. Deep Belief Network Structure

DBN is mainly divided into four steps in the process of training the model:

(1) Separately and unsupervised train the RBM network of each layer for data preprocessing, including noise removal, contrast enhancement and other operations. Ensure that the feature information is retained as much as possible when the feature vector is mapped to different feature spaces.

(2) Feature extraction: The network is designed to extract the features of the image and transform the original image into a high-dimensional feature vector.

(3) Feature selection: According to the importance and interrelationship of features, the most representative features are selected to reduce the feature dimension and improve the classification accuracy.

(4) Segmentation result generation: According to the classification results, the image is divided into different regions, and each region is processed, such as edge detection, morphological processing, etc.

Restricted Boltzmann machine is an energy-based model [18], which involves energy function and probability distribution. The energy function of the RBM model with a given state (v, h) is:

$$E(v,h|\theta) = -\sum_{i=1}^{n} a_i v_i - \sum_{j=1}^{n} b_j h_j - \sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij} h_j v_i \tag{4}$$

Where n represents the number of visible nodes, m represents the number of hidden nodes, w_{ij} represents the weight between the visible node i and the hidden node j, a_i represents the offset of the visual node, b_j is the bias of hidden nodes, v_i represents the state of the i th visible unit, h_j represents the state of the j th hidden unit.

According to the energy function, the joint probability of a visible node and a hidden node is defined, then the joint probability distribution is

$$\begin{cases} E(v,h|\theta) = \dfrac{1}{Z(\theta)} e^{-E(v,h|\theta)} \\ Z(\theta) = \sum_{v,h} e^{-E(v,h|\theta)} \end{cases} \tag{5}$$

Where $Z(\theta)$ denotes the normalization factor. $P(h_j = 1|v)$ and $P(v_i = 1|h)$ can be calculated from the joint probability.

$$\begin{cases} P(h_j = 1|v) = \sigma(b_j + \sum v_i w_{ij}) \\ P(v_i = 1|h) = \sigma(a_i + \sum h_j w_{ij}) \end{cases} \tag{6}$$

Where $\sigma(\cdot)$ function is the activation function of the neural network.

For a given training sample, the purpose of training RBM is to adjust the parameters to fit the given training data set. The maximum likelihood function derivative method is often used to solve the parameters [19]. The definition of the maximum likelihood function is

$$\theta^* = \arg\theta \max \sum_{t=1}^{T} \ln P(v^{(t)}|\theta) \tag{7}$$

Where T denotes the number of samples.

The DBN model has strong expansibility and flexibility. Many scholars have applied DBN to image segmentation. Traditional medical images are manually extracted features, which have certain subjectivity and lead to diagnostic errors. The emergence of DBN overcomes the shortcomings of manual extraction and can be used for image segmentation with high accuracy.

3. Scheme Design and Model Building

In order to compare the application effects of CNN and DBN in oral medical image diagnosis research, this study will build oral image analysis systems based on CNN and DBN respectively, and carry out comparative experiments. The specific experimental scheme block diagram is shown in Figure 4.First, the tooth slices were preprocessed to extract the disease area. Then, the CNN and DBN models are trained respectively, and the results are analyzed and evaluated. Finally, by comparing the experimental results of CNN and DBN models, the performance differences between the two deep learning methods in oral disease imaging diagnosis are evaluated.

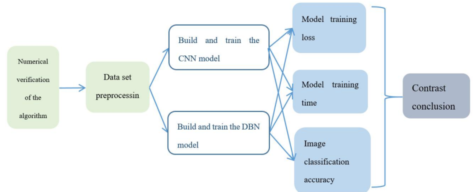


Figure 4. Experimental Scheme Block Diagram

3.1 Tooth Slice Data Set Construction

The data set used in this paper is CT images from the hospital, and the main case is dental jaw cyst.Fig.5 is a partial dataset image, and these datasets are named and numbered. Data preprocessing is performed on the test data set, including converting the selected data set into a grayscale image, reducing the frequency of pixel values, removing noise, and enhancing contrast. The purpose is to better identify the test image later, so that the subsequent program can be more easily compared.

The teeth of the data set were labeled by multiple professional imaging diagnosticians, and the lesion range was extracted according to the labeled information. Firstly, the pixels of the data set are processed in order to identify the test image later, and some data samples are selected as the test set (as shown in Table 1).

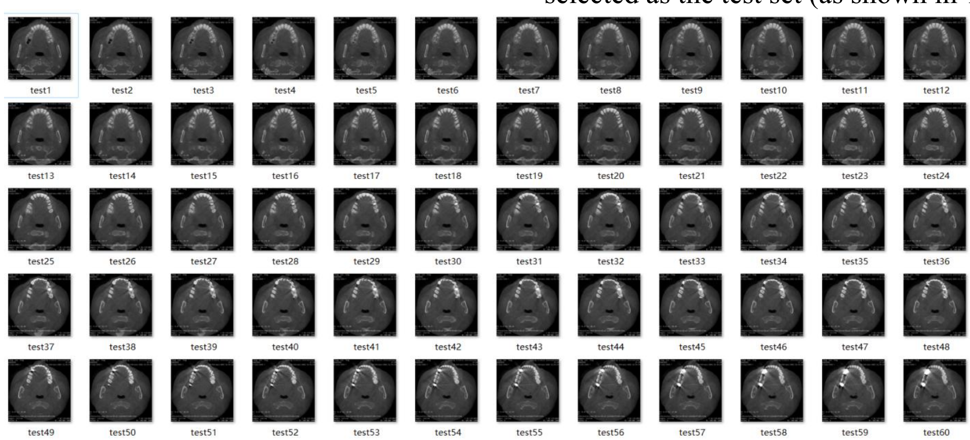


Figure 5. Part of the Dataset Image

Table 1. Number of Test Set Samples

	Number of frontal bone cyst samples	No jaw cyst sample number	grand total
testing set	45	15	60

3.2 CNN Experimental Design

3.2.1 Experimental process

Considering the basic characteristics of dental slices, the experimental process of oral disease imaging diagnosis based on CNN in this paper mainly includes three steps:

(1) According to the CT annotation data set given by the hospital, the experimental test set is selected for pixel processing.

(2) Design the corresponding network parameters to build the CNN model. Select the black and white pixels as the same as the teeth for training;

(3) According to the experimental results, the performance of the CNN model is evaluated from the parameters such as the accuracy of image classification, model training loss and training time.

3.2.2 Model construction

The experiment builds a 6-layer network structure, including 4 convolutional layers, 1 pooling layer and Softmax layer. The specific structure is shown in table 2.

Table 2. The Specific Structure of Convolutional Neural Network

tier number	type	size of convolution kernel	Output feature map size	Number of output feature maps
1	convolution layer	3×3	30×30	64
2	max pooling	2×2	15×15	64
3	convolution	3×3	13×13	128
4	transposed convolution	4×4	10×10	128
5	convolution	1×1	10×10	512
6	Softmax	-	-	2

The CNN training process was completed on a single GPU, and a semantic segmentation neural network was created and trained to extract features. The image was input into the built neural network and processed as follows:

(1) Considering the basic characteristics of the tooth slice, the system function is called, and the CT image to be tested is input into the basic network. After four convolutional layers (one of which replaces the fully connected layer), one pooling layer and softmax layer, the first training is performed, and the image is segmented and the test results are obtained.

(2) Improve the results, re-train and segment, in which the use of class weights to balance the class produces better segmentation results. Other steps to improve results include increasing the number of epochs used for training, adding more training data, or modifying the network.

3.2.3 Experimental results

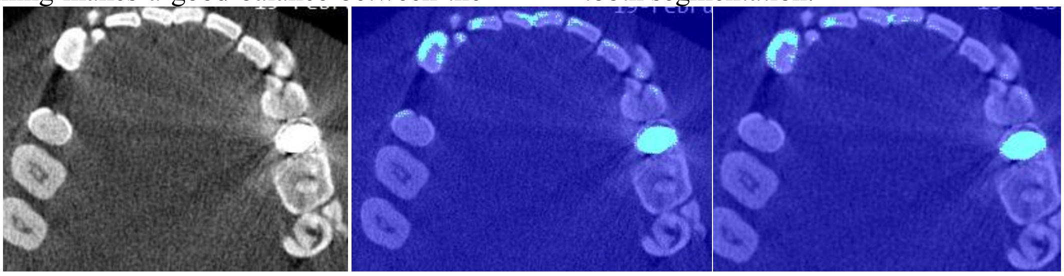
A total of two trainings were conducted in the experiment, and the small batch accuracy and small batch loss of the two trainings were compared. The results are shown in Table 3.

Table 3 Experimental Results of CNN Segmentation Network Training

	iteration times	Small batch accuracy	Small batch loss
First training	300	98.92%	0.0294
The second training	300	98.01%	0.0475

It can be seen from Table 3 that due to the relatively perfect structure of the CNN model, higher segmentation image accuracy can be obtained in the experiment, and the accuracy rate can reach about 98 %. It shows that the built model can accurately identify the location of the lesion, and it is not easy to misdetect the tooth segmentation.

The first training in the experiment is that the system trains itself in the training data. The second improvement is mainly to increase the total amount of data, so that it must include all the training data. The accuracy rate is slightly lower than the first training accuracy, but the accuracy is more than 95 %.The purpose of increasing the total amount of data is to add unnecessary pixel interference factors. Through training, the neural network can segment the test image more accurately and has stronger nonlinear fitting ability, so as to further improve the performance of the model. From the results of the program, it can be seen that the normal teeth of the two trainings are different. Figure 6 (a) is the original image of the test image, Figure (b) mislabeled the normal teeth, and Figure (c) almost only labeled the abnormal teeth. The improvement of training makes a good balance between the performance and computational cost of the model. The experimental data show that CNN has a high accuracy in segmenting images. The built model can accurately identify the location of the lesion and is not easy to misdetect in tooth segmentation.



(a)The Original Image (b) Images after the First Training (c) Images after the Second Training

Figure 6 Comparison of Two CNN Training Results

3.3 DBN Experimental Design

3.3.1 Experimental process

In this paper, the experimental process of image diagnosis of oral diseases based on DBN mainly includes four steps:

(1) Select the appropriate experimental test set according to the CT labeling data set given by the hospital;

(2) Preprocessing the test data set;

(3) Design the corresponding network parameters, build a DBN model, run apretrain, convert it into MAT format, input the data set, and transfer the program for feature extraction, feature selection and segmentation result generation;

3.3.2 Model building and experimental results

In the experiment, a four-layer DBN network structure is built, including one input explicit layer, two RBM invisible layers and one Softmax output layer, which are arranged from bottom to top. The specific structure is shown in Table 4.

DBN training process:

(1) First, fully train the first RBM;

(2) Fixing the weight and offset of the first RBM, and then using the state of its recessive neurons as the input vector of the second RBM;

(3) After fully training the second RBM, stack the second RBM above the first RBM;

(4) Repeat the above three steps for any number of times;

(5) If the data in the training set are labeled, then in the top-level RBM training, in addition to the dominant neurons, neurons representing classification labels are needed in the RBM display layer to train together.

Table 4 Concrete Structure of Deep Confidence Network

number of plies	Each layer structure	type	Number of nodes
one	RBM1	Visible layerv^1	2500
2	RBM2	Hidden layerh^1	1500
three	RBM3	Hidden layerh^2	500
four	Softmax	-	2

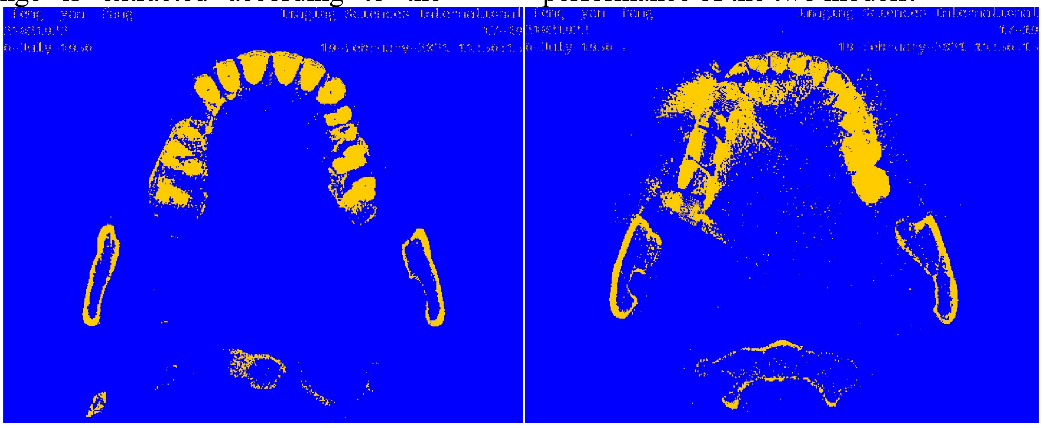
Among them, the number of nodes in the visible layer is the pixel value of the image. The hidden layer is the number of nodes that can extract the data related to the upper layer. In the Softmax layer, set the number of nodes to represent the output type.

In this section, a model of segmented image is built based on DBN, which is used to determine whether there are jaw cysts in the segmented image. Figure 7 (a) and (b) show the segmentation results of DBN images of two test samples, No.16 and No.57 respectively.

4. Comparative Analysis of CNN and DBN Experiments

CNN and DBN are both commonly used deep learning algorithms for image recognition and classification. Both algorithms can be used in the imaging diagnosis of oral diseases. Let's compare their performance from different dimensions of experimental data.

The data set includes jaw cyst images and jawless cyst images. The dental slices of the data set are marked by a number of professional image diagnostic doctors, and the lesion range is extracted according to the marking information, in which the highlighted part of the image is the tooth lesion. Data sets with obvious lesion characteristics are selected as test sets and input into CNN and DBN models respectively, and then the output image results are compared to evaluate the performance of the two models.

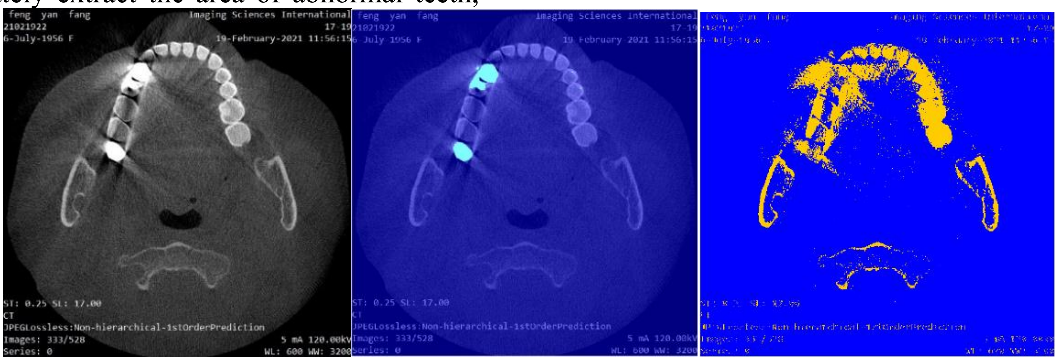


(a) Test16 (b) Test57

Figure 7 DBN Image Segmentation Results

4.1 Image Distortion Contrast

Taking the original image of Test57 as an example (as shown in Figure 8(a)), more astigmatism may cause a complex oral image environment. Fig. 8(b) shows CNN segmentation result, and fig. 8(c) shows DBN segmentation result. By comparing Figure 8(b) and Figure 8(c), it can be seen that CNN can accurately extract the area of abnormal teeth, while DBN is affected by complex environment, which leads to the great difference between the extracted area of abnormal teeth and the original image. Some useful information of two abnormal teeth located in the upper left area of the image has been eliminated, which leads to local area distortion of DBN processing.

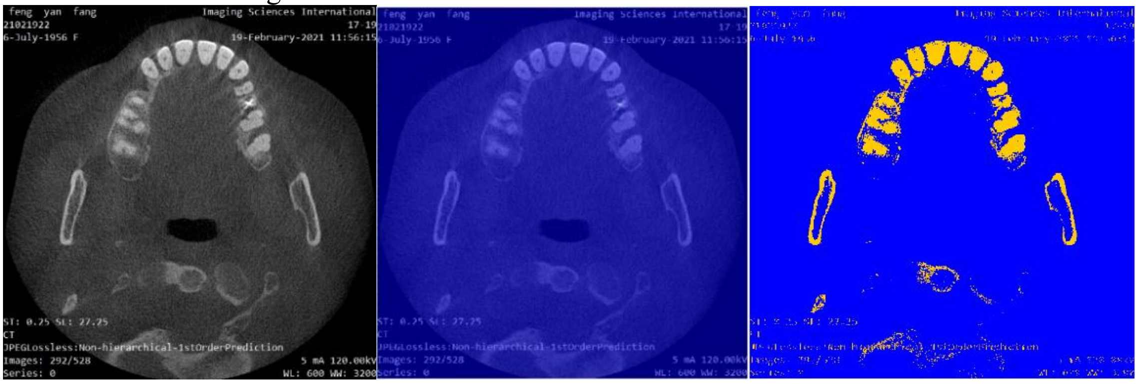


(a) Original Image (b) CNN Segmentation Result (c) DBN Segmentation Result.

Figure 8 Segmentation Results of Case No.57 Image

4.2 Comparison of Feature Extraction Accuracy

Taking case No.16 as an example, the results are shown in Figure 9.Among them, Figure (a) is the original image, Figure (b) is the CNN segmentation result, and Figure (c) is the DBN segmentation result. From the comparison results of this group, the distinction between normal teeth and abnormal teeth in CNN is obvious. For example, the center black area of each tooth in the figure and the boundary and contour of the tooth are clear. For the results of DBN, the characteristics of the central black region of some normal teeth were not extracted. Although the central black region of the right abnormal teeth was also more obviously extracted, the boundary contour of the teeth was different from the original image, while the boundary of the periodontal region of the CNN results was more clearly displayed. It shows that CNN has high sensitivity in detail processing of dental images. Because the two models have different image feature extraction methods, CNN can extract complex tooth image features more effectively.



(a) Original Image (b) CNN Segmentation Results (c) DBN Segmentation Results

Figure 9. Image Segmentation Results of Case No.16

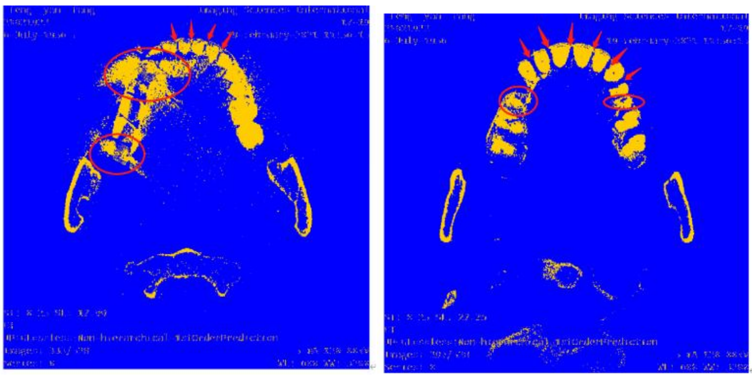
4.3 Comparison of Tooth Discrimination

The data set with obvious lesion features was selected as the test set and input into the built DBN model respectively. The lesion parts of the segmentation results of the DBN model were marked twice and judged. Take the comparison results of Test57 (shown in Figure 10 (a)) and Test16 (shown in Figure 10 (b)) as an example:

As shown in Fig.10 (a), the red circle is the clarity of the periodontal boundary. Normal teeth: the left yellow area, between the two red circles adjacent to the normal teeth and abnormal teeth is difficult to distinguish.

As shown in Fig.10 (b), the red circle is the periodontal boundary, which is blurred compared with the original image, and the black area in the center of the tooth is not obvious, which makes it difficult to distinguish between normal teeth and abnormal teeth.

The comparison results show that DBN can not detect the lesion site stably when segmenting the image, and there may be false detection in tooth segmentation. It has strong instability in the classification and recognition of diseased teeth and the error rate of diagnosis with or without diseased teeth.



(a) Test57 (b) Test16

Figure 10. Comparison of Tooth Discrimination

4.4 Accuracy Comparison

The model is adjusted, trained, evaluated on the validation data, and the model is adjusted again by trying different architectures and different hyperparameters, and then the process is repeated until the model achieves the best performance.

Different training times are selected to compare the accuracy of the two models in image segmentation, and the performance of the two models is evaluated. The experimental results are as Table 5.

Table 5. Comparison of the Results of Different Training Times and Test Set Accuracy

frequency of training / model	100	200	300
CNN model	97.97%	98.79%	98.78%
DBN model	96.01%	96.99%	95.33%

Although both CNN and DBN models have

good performance for image processing, due to the different feature extraction principles of these two models, the results of image segmentation performance are different. It can be seen from the data in Table 5 that with the increase of training times, the training accuracy of the two models is basically increasing, but regardless of the number of training times, the accuracy of the DBN model is always lower than that of CNN. The accuracy of CNN tends to be stable after training reaches a certain number of training times, and fluctuates in a small range around 98 %. DBN is not stable enough compared with CNN. It can be seen that the CNN model will be more stable.

4.5 Experimental Conclusion

By comparing the above experimental results, the following conclusions can be drawn:

(1) DBN is more suitable for processing unstructured data. It can automatically denoise the data and extract high-dimensional feature representations, which can be used to construct a high-precision oral disease diagnosis model.

(2) Comparing the performance of the two models from the above four different dimensions, it can be seen that CNN is more stable. Due to the different feature extraction principles of the two models, the obtained image segmentation performance results are different.

CNN and DBN are commonly used image segmentation algorithms. They have good performance in image segmentation. In practical applications, the two methods are usually selected according to different task requirements. From the comparative analysis of the above experiments, it can be concluded that the CNN model is more accurate in distinguishing normal teeth from abnormal teeth, the accuracy of abnormal tooth feature extraction in non-complex image environment, and image distortion (whether to distort the local useful information of tooth image). It performs better in the classification and recognition of diseased teeth, and has a low error rate for the diagnosis of diseased teeth. With the increase of the number of iterations, CNN will be more stable, while the recognition performance of DBN model needs to be further improved. DBN is an important model for processing images in deep learning. Compared with CNN, which has built-in functions in MATLAB that can be called directly, DBN does not, so the network structure is relatively simple. Moreover, DBN should require a large amount of data for training and learning. However, due to the lack of data sets, it may lead to problems such as over-fitting. It can only compare and analyze the advantages and disadvantages of the model from a large number of test image results. Therefore, finding more suitable data sets and setting more layers of the model may obtain more accurate experimental results, and further experiments will be carried out later to make further improvements.

5. Summary

In this paper, the data set of frontal bone cyst cases was constructed by selecting the CBCT image set of basic teeth cases. Through designing and building two neural network models of CNN and DBN, the image segmentation results of the case image are segmented from different dimensions. The main research work and conclusions of this paper reflect the good performance of CNN and DBN for image segmentation, and verify that the performance of CNN is better than that of DBN. However, there are still some improvements in the field of oral imaging detection in the future:

(1) Part of the code will be repeated when building the model, which needs to be further integrated and modified.

(2) When building a CNN model, most of them consider processing local information and summarizing overall information to adjust their network structure. These two aspects are inseparable and inseparable. When balancing the relationship between the two, the complexity of the model should be considered. The complexity of the CNN model is determined by the number of layers built, the number of convolution kernels, and the size of the convolution kernel. For the classification and segmentation of oral medical images, some algorithms will use manual labeling when locating features. Doctors only give corresponding judgment results based on the results of labeling, and will not mark the feature area in the teeth.

(3) In the CNN model, the original image is subjected to multiple convolutions and down-sampling until the output result is obtained. Therefore, the convolution output signal

directly determines the mapping relationship between input and output. The resolution of the original CT image is high. If the original CT image is directly input into the built model, the number of pooling will increase, resulting in too many model parameters. The configuration requirements of the computer will also increase, which may exceed the solving ability of the computer. Because the image features of medical images are less specif ic than the graphic features of general scenes, it is a feasible improvement direction to obtain the approximate area of features. In this experiment, the resolution of CT image is reduced to 640×640, which will lose some image features. If the feature region of the image can be obtained first, and then the larger resolution can be used to analyze the region with lesions, more accurate experimental results may be obtained, and experiments will continue to be carried out in the later stage to make further improvements.

In future research work, transfer learning can be considered. Transfer learning can solve the problems of insufficient data sets and over-fitting, and can make the network training and loss convergence faster to meet the requirements of efficient detection of various oral diseases in the future.

Acknowledgments

We thank for financial support from the National students' innovation and entrepreneurship training program of Tianjin University of Commerce (202310069012).

References

[1] Y. Hong and F. Lv, "Data Mining and Knowledge Discovery in Computer Aided Medical Diagnosis System," 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 2021, pp. 797-800.

[2] Woldeamanuel Minwuye Mesfin, Kim Taehoon, Cho Soojin, Kim Hyeong-Ki. Estimation of concre-testrength using thermography integrated with deep-learning-based image segmentation: Case studies and economic analysis. Expert Systems with Applications, 2023, 213(3): 119-249.

[3] WANG Yanlin, LI Gang. Research progress on artificial intelligence in imaging diagnosis of oral diseases. Journal of Prevention and Treatment for Stomatological Diseases, 2022, 30(11): 816-820.

[4] Kim I, Kang K, Song Y, Kim TJ. Application of Artificial Intelligence in Pathology: Trends and Challenges. Diagnostics (Basel). 2022 Nov 15; 12(11):2794.

[5] Zhao Q, Zhang B X, Lyu S C, et al. A CNN-SIFT Hybrid Pedestrian Navigation Method Based on First-Person Vision. Remote Sens, 2018, 10(8): 23.

[6] Kaur M, Singh D. Fusion of medical images using deep belief networks. Cluster Comput, 2020, 23(2): 1439-53.

[7] Tang B B, Liu X, Lei J, et al. DeepChart: Combining deep convolutional networks and deep belief networks in chart classification. Signal Process, 2016, 124: 156-161.

[8] Yang R Y, Singh S K, Tavakkoli M, et al. Cnn-lstm deep learning architecture for computer vision-based modal frequency detection. Mechanical Systems and Signal Processing, 2020, 144.

[9] Tuzoff DV, Tuzova LN, Bornstein MM, et al. Tooth detection and numbering in panoramic radiographs using convol l networks. Dentomaxillofac Radiol, 2019, 48(4): 51-61.

[10] Poedjiastoeti W, Suebnukarn S. Application of convolutional neural network in the diagnosis of jaw tumors. He Res, 2018, 24(3): 236-241.

[11] Ch oi J, Eun H, Kim C. Boosting Prox imal Dental Caries Detection via Combination of Variational Meth ods and Convolutional Neural Netw ork. J Signal Process Sys, 2018, 90(1): 87-97.

[12] Orh an K, Bayrakdar IS, Ezh ov M, et al. Evaluation of artificial intelligence for detecting periapical path osis on conebeam computed tomograph y scans. Int Endod J, 2020, 53(5): 680-689.

[13] Mohamed A R, Yu D, Deng L. Investigation of full-sequence training of deep belief networks for speech recognition. INTERSPEECH 2010, Conference of the International Speech Communication Association, Makuhari, Chiba, Japan, September. DBLP, 2010:2846-2849.

[14] Dahl G E, Yu D, Deng L, et al. Context-

Dependent Pre-Trained Deep Neural Networks for Large Vocabulary Speech Recognition. IEEE Transactions on Audio, Speech and Language Processing, 2012, 20(1): 30-42.

[15] Yang Chuanshu, LI Changsheng, SUN Xudong, HUANG Liming, ZHANG Haolin. Research Method and Practice of Artificial Intelligence Drilling Technology. Petroleum Drilling Techniques, 2021, 49(5): 7-13. doi: 10.11911/syztjs.2020136

[16] Geetha Chandrashekar, Saeed AlQarni, Erin Ealba Bumann, Yugyung Lee. Collaborative deep learning model for tooth segmentation and identification using panoramic radiographs. Computers in Biology and Medicine. 2022, 148, 105829.

[17] Lee K S, Jung S K, Ryu J J, et al. Evaluation of Transfer Learning with Deep Convolutional Neural Networks for Screening Osteoporosis in Dental Panoramic Radiographs. Journal of Clinical Medicine, 2020, 9(2):392.

[18] Zhang Z, Zhao J.A deep belief network based fault diagnosis model for complex chemical processes. 2017, 54(4): 395-407.

[19] Marlin B M, Swersky K, Chen B, et al. Inductive Principles for Restricted Boltzmann Machine Learning. journal of machine learning research − proceedings track for artificial intelligence & statistics, 2010.

