Data Driven Intelligent Diagnostics for Parkinson's Disease

YIN DAI [1,2], **ZHENG TANG** [1], **YANG WANG** [3,4], **AND ZI'AN XU** [1]
[1]College of Medicine and Biological Information Engineering, Northeastern University, Shenyang 110169, China
[2]Engineering Center on Medical Imaging and Intelligent Analysis, Ministry Education, Northeastern University, Shenyang 110169, China
[3]School of Computer Science and Engineering, Northeastern University, Shenyang 110169, China
[4]Neusoft Corporation, Northeastern University, Shenyang 110169, China
Corresponding author: Yin Dai (daiyin@bmie.neu.edu.cn)

ABSTRACT Parkinson's disease (PD) is a common chronic occult neurological degeneration disease which is most likely to occur among middle-aged and old-aged people. Neuroprotective methods can significantly slow the PD progression, therefore early diagnosis and the treatment of the PD is crucial. As the second most common neurodegenerative disorder, Parkinson's disease has its identification based on clinical diagnoses. Early diagnosis is especially important due to the lack of radical treatment. In all early diagnostic methods, results of imaging diagnosis based on positron emission tomography (PET) imaging has the most outstanding accuracy. Deep learning methods based on convolutional neural networks performed well in diagnosing different diseases. To address the demand of early diagnosis, this paper applies several preprocessing methods to image data such as gray level transformation, histogram equalization, improved wavelet soft-threshold denoising and image enhancement, and proposes a deep learning model based on U-Net architecture with deformable convolution kernels. After comparing the results of both methods, we found out that deformable U-Net outperformed the previously built improved VGG-Net. These results show that U-Net with deformable convolution structure has a good diagnostic capability and a better performance than the improved VGG-Net.

INDEX TERMS Parkinson's disease, PET, U-Net, VGG-Net, CNN, data analytics.

I. BACKGROUND

A. STATUS OF PARKINSON'S DISEASE

Parkinson's Disease (PD) is a common chronic occult neurological degeneration disease which is most likely to occur among middle-aged and old-aged people. Clinical symptoms of PD can be divided into motor symptoms, which include resting tremor, bradykinesia, postural instability, etc., and non-motor symptoms [1], which include mental symptom, cognitive function change, language barrier and autonomic nervous dysfunction [2]. Currently, most PD clinical symptoms are not obvious. It's difficult to make accurate diagnosis while only depending on its clinical manifestations and series routine examinations. When dopaminergic neurons in substantia nigra of brain decreases by 30% to 70%

than usual, the chance of early treatment would be already lost [3].

However, according to researches about PD, taking neuroprotective therapies before most of the dopaminergic neurons of the Substantia Nigra (SN) degenerates in the early stage of motor symptoms can help slow the disease progression [4]. Therefore, early diagnosis and early treatment of PD are particularly important. Speech-based data has been used as a material for diagnosing Parkinson's Disease, which applied series of deep learning algorithm to diagnose individual with the disease and the average clinical diagnosis accuracy reached 86.5% [1]. But audios samples contain a lot of noise and not able to extract visual features. Single-photon Emission Computed Tomography (SPECT) imaging is a reliable method for diagnosing Parkinsonism. ML algorithm was used for classification task, which achieved 71% to 82.2% of Health Controls vs PD binary classification. However,
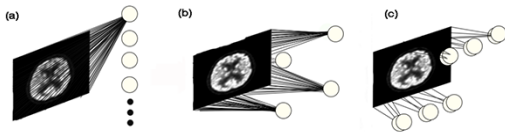
FIGURE 1. (a) A fully connected neural net example. If it is a 1000*1000 image with 1M hidden units, there will be 1012 parameters. (b) A locally connected neural net example. If it is a 1000*1000 image with 1M hidden units and its filter size is 10*10, there will be 108 parameters. (c) A convolutional neural net example. If it is a 1000*1000 image with 100 filters and the filter size is 10*10, there will be 104 parameters.

FIGURE 2. The flow chart of the preprocessing.

the extracted features used in this project usually misclassified SWEDD and PD patients [5]. To meet this issue head-on, the convolutional neural network is introduced to assist doctors to diagnose with PET image.

B. PET IMAGING DIAGNOSIS OF PARKINSON'S DISEASE

In the early stage, diagnosis of PD is mostly based on clinical diagnosis, and the rate of misdiagnosis is high. But with the development of medical imaging, imaging diagnosis with high accuracy of PD becomes important. PET images of PD patients are selected for computer-aided diagnosis in this study. Glucose uptake is the only way for brain to obtain energy, so difference of glucose uptake in various regions of the brain can show the extent of Parkinson's disease. If the biologically active molecule chosen for PET is an analogue of glucose, the concentrations of tracer imaged will indicate tissue metabolic activity as it corresponds to the regional glucose uptake [6].

^{18}F-FDG PET shows that patients with PD have a lower glycometabolism in frontal cortex and a higher one in parietal cortex, which is aggravated as time prolonged [7]. PET shows glucose metabolism changes in the patient's brain, which can educe distribution of the lesion.

C. DEEP LEARNING AND CONVOLUTIONAL NEURAL NETWORK

With the development of cloud computing, study of complex learning models represented by deep learning becomes popular in the early 21st century. Deep learning can be interpreted in a narrow sense as a multi-layer neural network. Although the model is complex, computing power can be significantly improved and the risk of overfitting can also be reduced when the data size is huge. It lowers the barrier of machine learning and brings convenience for machine learning to engineering practice. Recently, deep learning has been a hot research topic in machine learning, and it has contributed in the fields of speech recognition and computer vision [8]–[10].

In machine learning, a Convolutional Neural Network (CNN) is a class of deep, feed-forward artificial neural networks that has successfully been applied to analyzing visual imagery. CNN uses a variation of multilayer perceptron designed to have minimal preprocessing requirements. Figure 1 provides a schematic overview of the method used for reducing parameter of neural network.
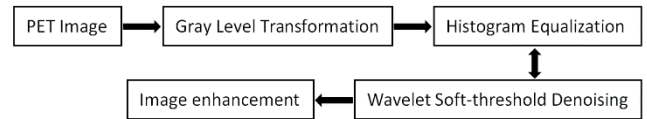
Deep learning methods using convolutional neural networks have achieved great success in the field of medical image analysis [11]. A special type of CNN is the U-Net [12], which is an example of image segmentation, for which labels were assigned to each pixel in the input image. There has been a lot of examples using U-Net for radiomic data analysis. In this paper, we applied U-net to identify PET images data and diagnose PD. There are examples of deep-learning aided PD diagnosis, but the methods are to analyze body-sensor data [13]. To the best of our knowledge, this is the first approach of applying deep learning method analyzing PET image data to diagnose PD.

II. METHODS

In this study, gray level transformation, histogram equalization, improved wavelet soft-threshold denoising and image enhancement were used for image preprocessing, then the convolutional neural network was trained and tested according to the PET image sets collected from Parkinson patients.

A. SOURCE OF PET IMAGES

A large amount of data is needed for training convolutional neural networks. To meeting such demand, 214 PET images with Parkinson's disease and 127 PET images of control group were collected from Parkinson's disease database – PPMI database, Alzheimer's disease database – ADNI database, and Human Connectome Project (HCP) database.

B. IMAGE DATA PREPROCESSING

Image data was collected from the open database. In order to ensure the image provided to train the convolutional with a high degree of recognition, image preprocessing is necessary before training. Schematic map of the preprocessing is shown in Fig. 2.

1) GRAY LEVEL TRANSFORMATION AND HISTOGRAM EQUALIZATION

The collected medical image is DICOM format. The gray value of DICOM format is −2000 to 2000. It is necessary to map the gray level of medical image to 0 to 255. After linear gray level transformation, the histogram equalization of the image is also needed. Histogram equalization is an important part of medical image preprocessing. Using histogram equalization for contrast adjustment, local contrast of the original image can be increased by equalizing brightness distribution, which can compensate the gray level in original image which is difficult to distinguish visually [14]. Histogram equalization can be expressed by the following
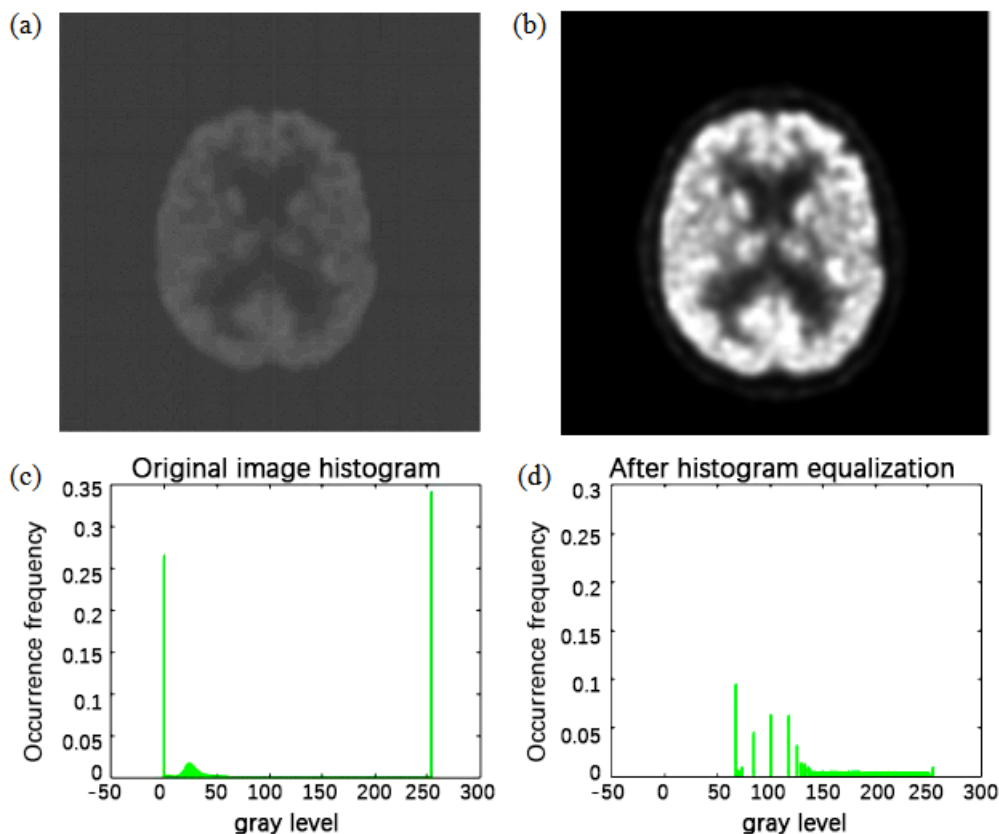
FIGURE 3. (a) The original image. (b) The image after gray level transformation and histogram equalization. (c) Original image histogram. The abscissa is gray level, and the ordinate is the frequency of occurrence. (d) Histogram after the histogram equalization of the image. The abscissa is gray level, and the ordinate is the frequency of occurrence.

formula:

$$s_k = T(r_k) = (L-1)\sum_{j=0}^{k} p_r(r_j) = \frac{(L-1)}{MN}\sum_{j=0}^{k} n_j \quad (1)$$

where L is the number of possible gray levels in the image, k is the gray scale level to be equalized, M is the length of image, N is the width of image, MN is the total number of pixels in the image, and n_k is the number of pixels in the r_k gray scale.

Take a PET image from PPMI database as an example, Fig. 3 shows the effect of gray level transformation histogram equalization.

2) WAVELET DENOISING
Image is susceptible to noise during the process of generation and propagation, which brings inconvenience to the diagnosis for doctor. It is also crucial to remove noise during image preprocessing. In general, image denoising based on frequency domain mainly uses low-pass filter, which can remove noise of high frequency [15]. The core of frequency domain wavelet denoising algorithm is multiresolution resolution, which can effectively remove the Gauss white noise [16]. It is superior to other denoising algorithms, but leads to the weakening of

medical image features. Here is an improved wavelet denoising method which is used in this study.

This denoising algorithm is based on the wavelet soft-threshold algorithm. When absolute value of the coefficient x is smaller than the global threshold α, the coefficient is assigned to 0 values, and the coefficient with an absolute value is made greater than the global threshold tends to zero, this is the soft threshold algorithm:

$$f(x) = \begin{cases} sign(x) \times (|x| - \alpha), & |x| \geq \alpha \\ 0, & |x| \leq \alpha \end{cases} \quad (2)$$

where the global threshold is calculated from the signal dispersion number N and the standard deviation of the Gaussian white noise by formula (3):

$$\alpha = \sigma^2 \sqrt{2\log(N^2)} \quad (3)$$

The threshold calculated by this method is larger and may cause details missing. Because wavelet coefficients are similar, here proposes a new method to calculate the threshold. Take square sum of the coefficients T in the window centered on wavelet coefficient, and calculate corresponding shrinkage
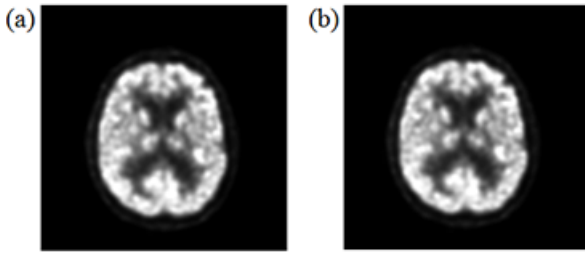
FIGURE 4. (a) Image before the improved wavelet denoising. (b) Image after the improved wavelet denoising.

FIGURE 5. Contrast stretching rules. Pixels in different gray level range with different transformation rules. Original range 0-80 linear mapping to 0-40, 80-175 mapping to 40-210, 175-255 mapping to 210-255.

FIGURE 6. (a) Image before contrast stretching. (b) Image after contrast stretching.

factor θ according to formula (4):

$$f(x) = \begin{cases} 1 - \dfrac{\alpha^2}{T^2}, & 1 \geq \dfrac{\alpha^2}{T^2} \\ 0, & 1 < \dfrac{\alpha^2}{T^2} \end{cases} \tag{4}$$

According to formula (5), find the global threshold and determine whether the wavelet coefficient tends to 0 shrink or set 0:

$$\alpha = \theta \times \sigma^2 \sqrt{2\log(N^2)} \tag{5}$$

The improved algorithm processed image is shown in Fig. 4.

3) IMAGE ENHANCEMENT

This work implements the contrast linear stretch transformation function for image enhancement, in which the input function and the output function of the gray-scale transformation is linear:

$$r = f(r) = k \times r + b \tag{6}$$

r is the gray level of the image. By changing parameter k and b, the gray transformation can be different:

(a) if $k > 1$, the contrast will be increased; if $k < 1$, the contrast will be decreased.

(b) if $k = 1$, $b = 0$, keep the original image; if $k = -1$, $b = 255$, the original image will be reversed.

In this paper, three stage contrast stretching is adopted. As shown in Fig. 5, the contrast stretching rules of pixels in different gray level ranges are different. Fig. 6 shows a schematic diagram of image before and after contrast stretching.

C. COMPUTER-AIDED DIAGNOSIS BASED ON CONVOLUTIONAL NEURAL NETWORK

1) STRUCTURE OF DEFORMABLE U-NET

U-Net is a special form of conventional CNN, which extends the typical CNN contracting paths in a symmetric way. Extracted information from the pooling layers are propagated to higher layers by the operations of up-sampling. Combining down-sampling ways with up-sampling ways present a U-shape. However, as traditional convolution method, U-Net has the shortcoming of processing object shape transformation due to its fixed perceptron. Inspired by the deformable
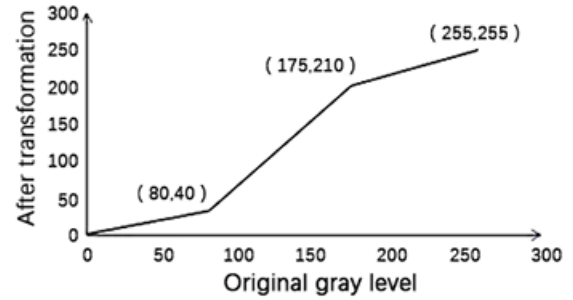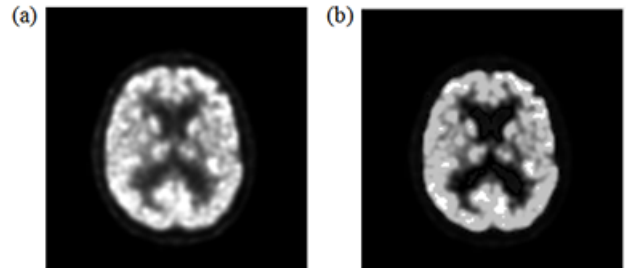
convolution neural network which was published by Dai in 2017 [17], this work replaces typical convolution kernel with deformable convolution in the whole network. The improved U-Net architecture is shown in Fig.7.

This U-Net, while similar to the original U-Net, has deformable convolution kernels. The network is organized with an encoder path and a decoder path which is formed by three layers. In each layer of encoder path, there exists two 3×3 deformable convolution kernels and a max-pooling operator with the step length of 2 for down-sampling. There is the base layer after the encoder layer, which is formed by two 3×3 deformable convolutions.

Similarly, each decoder layer consists of two 3×3 deformable convolution kernels and a max-pooling operation. The copy and crop connection between encoder and decoder can preserve more continuous information for localization [10]. The activation function we used in this network is ReLU. Network is optimized by the Adam-Optimizer, a method based on gradient decent, which makes the value of parameters relatively stable and promotes the dynamic adjustment of parameters.

Deformable convolution kernel is defined as:

$$Z^{l+1}(p) = \sum_{k=1}^{K1} \sum_{x=1}^{f} \sum_{y=1}^{f} [Z_k^l(p_0) w_k^{l+1}(\Delta p_n)] \tag{7}$$

where Z^l and Z^{l+1} are the input and output of convolution layer l, which is also called as feature map, $Z(i, j)$ denotes the value on pixels of the feature map. f, s_0, $K1$ each
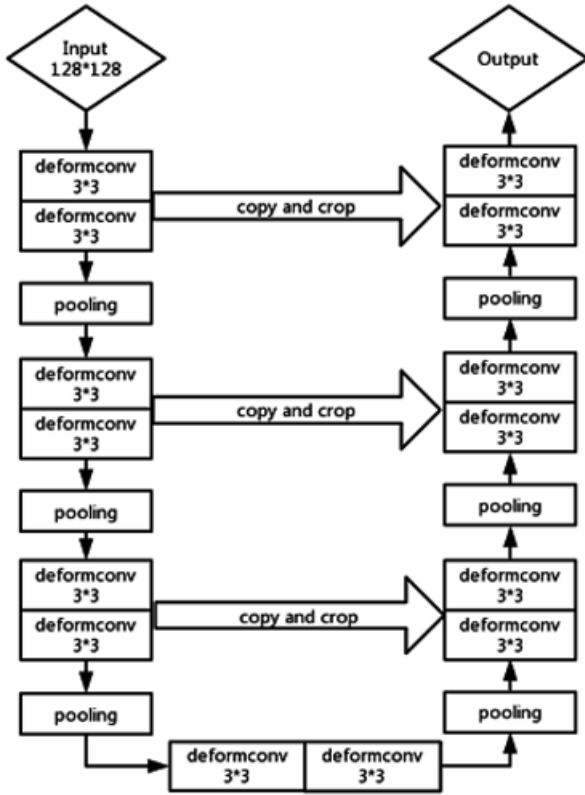
FIGURE 7. Architecture of the U-Net in this work.

represents size of convolution kernel, stride of kernel, and size of padding layers.

Here Δx Δy can be defined as fraction, so w_k^{l+1} can be bilinear interpolated and be expressed in the following form:

$$w_k^{l+1}(p) = \sum_{x=1}^{q} \sum_{y=1}^{q} u(q_w, p_w) u(q_z, p_z) w(q) \qquad (8)$$

where p stands for the arbitrary fractional pixels and q stands for the integral pixels in feature map. And u kernel can be defined as:

$$u(x, y) = max(0, 1 - |m - n|) \qquad (9)$$

This paper applies deformable kernels work in the same way throughout every layer, instead of implement different kernels for each layer to improve the performance of model. This network is trained end-to-end using the back-propagation method introduced in Table1.

2) TRAINING OF THE NETWORKS
Preprocessed Images labeled with PD and not PD are used to train the network.

The energy function used in this work is feature map pixels soft-max classifier combined with cross entropy loss function. Soft-max classifier is defined as:

$$s_k(x, y) = e^{A_k(x,y)} / \sum_{x=1}^{K} \sum_{y=1}^{K} e^{A_k(x,y)} \qquad (10)$$

TABLE 1. The steps of the back propagation algorithm.

Back Propagation algorithm
1. Initialization the network parameters.
2. Input a sample X and its corresponding expectation output Y.
3. Calculated the output of each layer, and the output of the i neuron on the k layer is: $U_i^k = \sum_{j=1}^{n+1} w_{ij} X_j^{k-1}, (X_{n+1}^{k-1} = 1, W_{i,n+1} = -\theta), X_i^k = f(u_i^k)$
4. According to the error cost function, the error of each layer parameter is found, for the output layer, $k = m$; $d_i^m = X_i^m(1 - X_i^m)(X_i^m - Y_i)$ Other layers: $d_i^k = x_i^k(1 - x_i^k) \sum_l W_{li} d_l^{k+1}$
5. Update the weight coefficient and threshold: $W_{ij}(t+1) = W_{ij}(t) - \eta d_i^k X_j^{k-1}$ $W_{ij}(t+1) = W_{ij}(t) - \eta d_i^k X_j^{k-1} + \alpha \Delta W_{ij}(t)$ Where $-\Delta W_{ij}(t) = -\eta d_i^k X_j^{k-1} + \alpha \Delta W_{ij}(t) = Wij(t) - W_{ij}(t-1)$

where $A_k(x, y)$ stands for the activation of the feature pixels in layer k, K is the number of pixels number in the feature map.

Then energy function can be expressed as follows:

$$E = \sum_{x=1}^{K} \sum_{y=1}^{K} \varpi(x, y) \log(s_{l(x,y)}(x, y)) \qquad (11)$$

where $\omega(x, y)$ is a predefined pixel weight map function, which gives each pixel a weight, in this paper we used the one given by Ronnenberger [12].

$$\varpi(x, y) = w_c(x, y) + \varpi_0 \exp(-\frac{(d_1(x, y) + d_2(x, y))^2}{2\sigma^2}) \qquad (12)$$

Initialization of the weight is important for the U-net, otherwise some pixels of feature map might give excessive activation which affects other channels and layers. According to the experiment, we set $\sigma = 8$ and $\varpi_0 = 10$ to have better performance. Epoch is set 10000, learning rate 0.01.

3) STRUCTURE OF IMPROVED VGG-NET
VGG-Net [18] is a convolutional neural network for image classification and recognition. In 2014, it achieved the first place of classification projects, and had good feature extraction ability. But VGG-Net involves many parameters, it is difficult to meet the demand of memory. This article improved the classic VGG-Net model, make the network more suitable for limited memory. The convolutional neural network model is shown in Fig. 8.

This network model is divided into three units, each consists of two convolution layers and one down-sampling layer. Size of convolution kernel is 5×5, and sliding step is 1. The stacking of the convolutional layer is to perform closer convolution operations on the input, so it can extract more features. The stacking of two 5×5 size convolution kernels equivalent to a 9×9 convolution kernel, but the stacked convolution layer has better nonlinear transformation performance and better feature extraction capability. Connect a fully connected layer and SoftMax classifier in the end and output the result
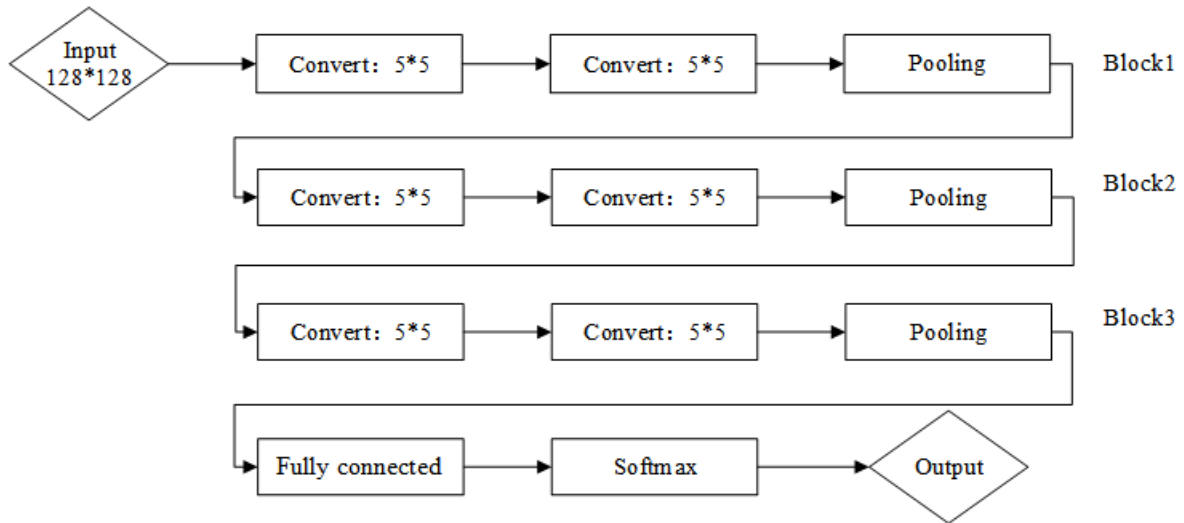
FIGURE 8. Schematic diagram of convolutional neural network.

of two-category. The activation function used after each convolution layer and fully connected layer is ReLU.

4) TRAINING OF IMPROVED VGG-Net
(1) Related Problems of Convolution Neural Network Algorithm

a: PARAMETER INITIALIZATION

Before the CNN model is trained, parameter initialization is momentous in the gradient descent algorithm. If initialization value is in the error surface which is relatively flat, the convergence rate will be slow, and the distribution of the initial weight is set as follows:

$$W \sim U\left[-\sqrt{6}/\sqrt{p^{(l)} + p^{(l-1)}}, \; \sqrt{6}/\sqrt{p^{(l)} + p^{(l-1)}}\right] \quad (13)$$

b: REGULARIZATION STRATEGY

In the convolutional neural network, network model with hidden layer can learn various nonlinear mapping relationships. The more layers there are, the stronger learning ability network will have. But lacking of data samples will cause overfitting.

In order to avoid this ill-posed problem, the regularization strategy is adopted. The description of network complexity is added into the cost function, amplitude of the weight is reduced, and expression of the error function after regularization is:

$$E^N = \lambda \frac{1}{2} \sum_{n=1}^{N} \sum_{k=1}^{c} (t_k^n - x_k^n)^2 + (1 - \lambda) \sum_i \omega_i^2 \quad (14)$$

where ω_i represents the connection weight and threshold, $\lambda \in (0, 1)$. That can be used to compromise both empirical errors and network complexity.

c: SoftMax CLASSIFIER

The value of the output layer of the convolutional neural network represents the classification results, so a classifier needs to be connected in the convolutional neural network, and the SoftMax classifier is used in this paper.

SoftMax classifier calculates the probability of each category according to the input of each sample, expression of the hypothetical function is as follows:

$$h_\theta(x) = \begin{bmatrix} p(y^{(i)} = 1|x^{(i)}; \theta) \\ p(y^{(i)} = 2|x^{(i)}; \theta) \\ \cdots \\ p(y^{(i)} = k|x^{(i)}; \theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^{k} e^{\theta_j^T x^{(i)}}} \begin{bmatrix} e^{\theta_1^T x^{(i)}} \\ e^{\theta_2^T x^{(i)}} \\ \cdots \\ e^{\theta_k^T x^{(i)}} \end{bmatrix} \quad (15)$$

where k represents the number of categories and θ is the general term for network parameters, $p(y^{(i)} = j|x^{(i)})$ represents the probability of the i-th sample $x^{(i)}$ input belongs to category j.

According to the cost function, parameter gradient of the convolutional neural network is obtained by iterative optimization algorithm, and weight value is updated in gradient descent strategy.

In order to prevent the parameter from being too large, regularization strategy is introduced to make cost function become convex function, and global optimal solution can be obtained by the gradient descent method. By minimizing the cost of error function, a SoftMax classifier can be trained and obtained.

5) CONVOLUTION NEURAL NETWORK PERFORMANCE TEST BASED ON MNIST DATABASE
Before training the network for disease diagnosis, it is necessary to test performance of the deformable U-Net and the CNN model that was already built. 60000 handwritten

TABLE 2. U-Net network accuracy measured by 10-fold cross-validation method. K is the number of tests.

K	Accuracy	K	Accuracy
1	92.35%	6	97.52%
2	94.12%	7	97.76%
3	95.62%	8	97.83%
4	95.98%	9	97.96%
5	96.88%	10	98.03%

TABLE 3. CNN network accuracy measured by 10-fold cross-validation method. K is the number of tests.

K	Accuracy	K	Accuracy
1	92.78%	6	97.74%
2	95.32%	7	97.88%
3	96.52%	8	97.92%
4	96.97%	9	98.01%
5	97.45%	10	98.05%

digital image set downloaded from the MNIST database was used to verify the performance of the constructed network. 10-fold cross-validation method is used here to evaluate the performance and generalization capability of convolutional neural networks.

10-fold cross-validation method is originated from k-fold cross-validation method, the key idea is to divide the data set D by stratified sampling to get k mutually exclusive subsets of similar size, and then train and text the network k times. This method ensures the mutual exclusion between train set and test set of network model and evaluate network performance more accurately. In order to ensure the consistency of data distribution of each subset, the stratified sampling formula adopted in this project is as follows:

$$D = D_1 \cup D_2 \cup \cdots \cup D_k, D_i \cap D_j = \varnothing, \quad (i \neq j) \quad (16)$$

In this project, data set downloaded from the MNIST database is divided by stratified sampling to get 10 mutually exclusive subsets of similar size, then train and test the network 10 times. Test results are shown in table 2&3.

The average accuracy measured by 10-fold cross-validation method is 96.864%. Obviously, with the increase of training times, the accuracy of the convolutional neural network is gradually increased, and at the last time accuracy rate reached 98%.

U-Net and CNN designed for this project can be used to recognize the image set by data training. Accuracy of the test of these network models is high, which means the network has good generalization ability, proving that the performance of the convolutional neural network and the setting of the hidden layer parameters are correct and the recognition ability can be improved by training. Difference of performance between these two networks is not significant in this test.

6) COMPUTER-AIDED DIAGNOSIS BASED ON CONVOLUTIONAL NEURAL NETWORK

PET images from PPMI and other databases were collected and processed. There are 214 sets of PET images with
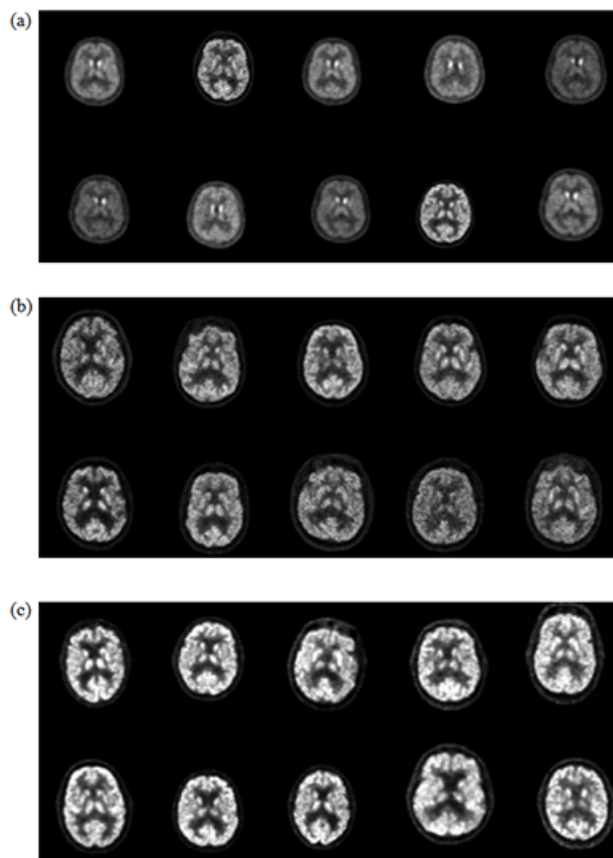

FIGURE 9. (a) Some images of patients with Parkinson's disease. (b) Partial sample of healthy subjects. (c) Some image of the image set after image preprocessing. All the PET images are processed before training.

Parkinson's disease (A patient's PET image is reserved for testing network diagnostic results) and 127 sets of PET images of healthy people. Figure 9 shows parts of samples and treatment.

Preprocessed image set is stored in a three-dimensional matrix. The matrix and its corresponding two classification labels are stored in the mat file, then read into the network. Because of the small amount of training and testing data, Hold-out method is used to verify the generalization ability of CNN.

Hold-out method divides data set D into two mutually exclusive sets, one is the training set S and the other one is test set T, the formula is expressed as $D = S \cup T$ and $S \cap T = \varnothing$. After training the model with S set, T set is used to test the model [19].

Finally, the reserved PET images of Parkinson's disease that are not used for training are read into the CNN model and diagnosed.

III. RESULTS

Using hold-out method, 253 images were used as training sets, and 88 images were the test set. After training and testing, diagnostic accuracy of the U-Net was 84.17% and the
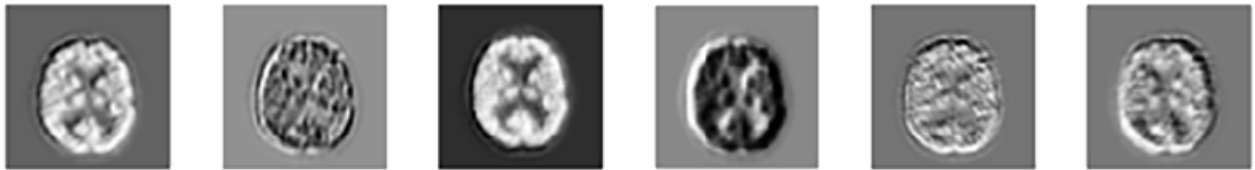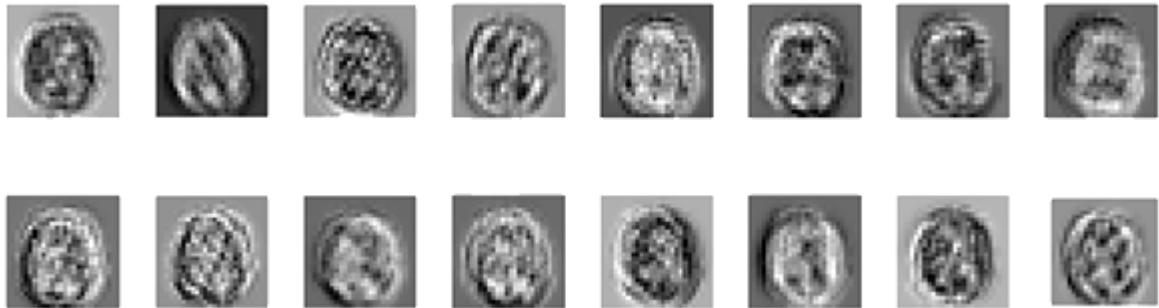
FIGURE 10. Output feature map of the first unit.

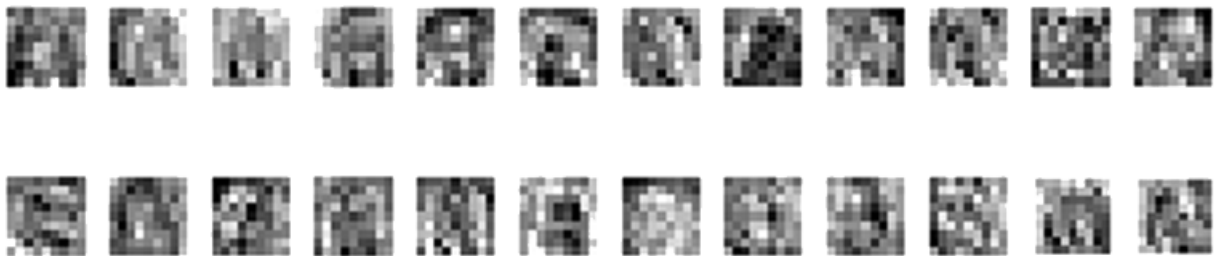**FIGURE 11.** Output feature map of the second unit. Size of map is 26 × 26.

FIGURE 12. Output feature map of the third unit. Size of map is 9 × 9.

CNN network was 76.19%, and the reserved PET images got correct diagnostic results.

According to the test results of neural network models, it can be concluded that U-Net performs better than the CNN in the task of PD diagnosis using PET images. To make trained convolutional neural network model perform more accurate disease diagnosis results, the network can automatically carry out characteristic learning, adjust weight and offset parameters during the diagnosis of Parkinson's disease. Convolutional neural network models applied in this work can extract and learn the characteristic information of Parkinson's disease image, then diagnose the disease more accurately.

IV. DISCUSSION

According to the computer-aided diagnostic results, convolutional neural networks can offer precise classification and accurate diagnosis of Parkinson's disease. The applied model is robust to the task of PET image data processing. U-Net with deformable kernels can deal with different kinds of noise in the PET images. The improved VGG-Net fits the limit memory working environment well and offers easy manipulating codes for users. Both VGG-Net and U-Net implemented in this work has the acceptable time cost. In the future, the larger size of samples would be applied to train this model, in order

to improve diagnostic performance and applied in the actual medical diagnosis.

Besides the output diagnostic result, the models in this work can also provide visualize feature map for doctors' further research or diagnostic. Principle of convolutional neural network is to learn features layer by layer, and extract image features by training the adjusting parameters. This project attempts to extract the feature map used to determine whether a patient is PD from the network model layer by layer, and make the feature map visible.

Fig. 10 shows the output feature map of the input image passed through the first unit. It can be observed from Fig. 10, in general, convolution layer of the first unit extracts some low-level features of basic grayscale and edge information.

Convoluted the input image with two convolution layer by trained convolution kernel, the output feature map can be obtained after down-sampling. Six feature map are printed, the size of a map is 60 × 60.

Fig. 11 shows the output feature map of the second unit. In the second stage, extraction of the convolution layer is more complex, and the convolutional neural networks learn features of some textures, which obtains more detailed than the previous stage.

Fig. 12 shows the output feature map of the Third unit. In the third stage, extraction of the convolution layer is more

discriminative, and each output feature graph is different. This is the discriminative feature learned by convolutional neural networks.

In summary, the feature learned by convolutional neural networks is discriminative enough. Visual feature maps extracted from the network can be fed back to clinicians, and it is of great value in clinical diagnosis research.

V. CONCLUSION

Due to the importance of early diagnosis and early treatment of Parkinson's disease, in this subject, the convolution neural network is introduced to assist doctors to diagnose PD with PET images. Firstly, preprocessed PET images collected from PPMI, ADNI and other databases. And then design and construct the networks, train and test it with the preprocessed image set. After training, the network models constructed to diagnose Parkinson's disease with accuracy rate of 84% and 75%. Meanwhile, the feature map can be visualized by convolutional neural network, and the feature of network learning can be extracted, which may provide further assistance to doctors in the diagnosis of Parkinson disease.

AUTHOR CONTRIBUTIONS

Dr. Yin Dai designed the algorithm. The project was conducted by Zheng Tang. Zheng Tang and Yang Wang analyzed the results, Yang Wang and Zi'an Xu prepared figures and tables and wrote the manuscript. All authors reviewed and approved the final manuscript.

COMPETING INTERESTS

The authors declare that they have no competing interests.

DATA AVAILABILITY

PET images are collected from Parkinson's disease database – PPMI database, Alzheimer's disease database – ADNI database, and Human Connectome Project (HCP) database. A primary goal of PPMI is to develop a large and comprehensive set of correlated clinical data and biospecimens that will be made available to the entire scientific community to help accelerate biomarker verification research.

The Alzheimer's Disease Neuroimaging Initiative (ADNI) unites researchers with study data as they work to define the progression of Alzheimer's disease. ADNI researchers collect, validate and utilize data such as MRI and PET images, genetics, cognitive tests, CSF and blood biomarkers as predictors for the disease. Data from the North American ADNI's study participants, including Alzheimer's disease patients, mild cognitive impairment subjects and elderly controls, are available from this site.

The purpose of this site is to build a centralized store of HCP-style protocols, which can be used to gather data from multiple populations with the purpose of comparing against the 1200 healthy young adults of the WU-Minn Human Connectome Project.

REFERENCES

[1] S.-I. Yang and Y.-C. Wu, "Progress of clinical research of Parkinson's disease dementia," *Chin. J. Contemp. Neurol. Neurosurgery*, vol. 17, no. 2, pp. 95–100, 2017.

[2] J.-Y. Zhang, P. Chan, and E.-H. Xu, "Attention must be paid to cognitive dysfunction in Parkinson's disease," *Chin. J. Contemp. Neurol. Neurosurgery*, vol. 17, no. 1, pp. 9–13, 2017.

[3] Y. Sun, X. Qi, and C. Li, "Progress of imaging examination in diagnosis of Parkinson disease," *Practical J. Cardiac Cerebral Pneumal Vascular Disease*, no. 7, pp. 8–11+16, 2014.

[4] Y. Li, "Classification of Parkinson's disease by decision tree based instance selection and ensemble learning algorithms," *J. Med. Imag. Health Inform.*, vol. 7, no. 2, pp. 444–452, 2017.

[5] D. Berg, "Transcranial ultrasound as a risk marker for Parkinson's disease," *Movement Disorders*, vol. 24, no. S2, pp. S677–S683, 2009.

[6] Y. Cheng, Z. Sha, and Y. Tian, "Imaging diagnosis of Parkinson's disease," *Neural Injury Funct. Reconstruct.*, no. 2, pp. 153–154+168, 2016.

[7] R. Juh, J. Kim, D. Moon, B. Choe, and T. Suh, "Different metabolic patterns analysis of Parkinsonism on the ^{18}F-FDG PET," *Eur. J. Radiol.*, vol. 51, no. 3, pp. 223–233, Sep. 2004.

[8] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Proc. Eur. Conf. Comput. Vis.*, vol. 8689, 2014, pp. 818–833.

[9] Y. Sun, X. Wang, and X. Tang, "Deep learning face representation from predicting 10,000 classes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 27, Jun. 2014, pp. 23–28.

[10] C. Farabet, C. Couprie, L. Najman, and Y. LeCun, "Learning hierarchical features for scene labeling," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1915–1929, Aug. 2013.

[11] D. Shen, G. Wu, and H. I. Suk, "Deep learning in medical image analysis," *Annu. Rev. Biomed. Eng.*, vol. 19, pp. 221–248, Jun. 2017.

[12] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2015, pp. 234–241.

[13] C. Rafael, E. Richard, and L. Steven, *Digital Image Processing*. Beijing, China: Publishing House of Electronics Industry, 2011.

[14] M. Zhang, X. Li, M. Xu, and Q. Li, "Image segmentation and classification for sickle cell disease using deformable U-net," 2017, *arXiv:1710.08149*. [Online]. Available: https://arxiv.org/abs/1710.08149

[15] W. Zhang, "The basic fundamental of X-CT and its application in medicine," *Sci. Technol. Inf.*, no. 30, pp. 188–193, 2010.

[16] K. Dong and S. Zhang, "Research on denoising method of magnetic resonance image based on wavelet analysis," *J. Chengdu Electromech. College*, no. 3, pp. 1–5, 2007.

[17] J. Dai, H. Qi, Y. Xiong, Y. Li, G. Zhang, H. Hu, and Y. Wei, "Deformable convolutional networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 764–773.

[18] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *Comput. Sci.*, 2014.

[19] M. Hassoun, *Fundamentals of Artificial Neural Networks*. Cambridge, MA, USA: MIT Press, 1995.

YIN DAI received the Ph.D. degree from the Department of Computer Science, Northeastern University, where she is currently a Lecturer with the College of Medicine and Biological Information Engineering. Her research interests include computer-aided diagnosis and medical image processing.
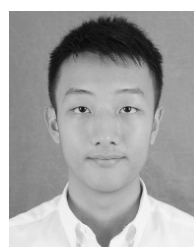
ZHENG TANG is currently pursuing the bachelor's degree with Northeastern University, where he is also doing research with the College of Medicine and Biological Information Engineering. His research interests include deep learning models, computer-aided diagnosis, and medical image processing.

YANG WANG is currently pursuing the Ph.D. degree from the School of Computer Science and Engineering, Northeastern University. His research interests include image processing and deep learning.

ZI'AN XU is currently pursuing the bachelor's degree with the College of Medicine and Biological Information Engineering, Northeastern University, where he is currently doing research with the College of Medicine and Biological Information Engineering. He is currently involved in medical images processing and the construction of deep learning models.

• • •