AUTOMATIC DIAGNOSTIC SYSTEM FOR PARKINSON'S DISEASE BASED ON DEEP LEARNING USING MIDBRAIN MAGNETIC RESONANCE IMAGES

DongYue Wang¹ and TaegKeun Whangbo^{2*}

¹*Dept. of Computer Science, Gachon University Sujeong-Gu, Seongnam-Si, Gyeonggi-Do, Korea*

²*Dept. of Computer Science, Gachon University Sujeong-Gu, Seongnam-Si, Gyeonggi-Do, Korea*

*tkwhangbo@gachon.ac.kr

Abstract— Magnetic resonance (MR) imaging of the midbrain is the primary means for the diagnosis of Parkinson's disease (PD). However, it is difficult to diagnose PD manually using MR images. Therefore, we developed an automatic diagnostic system for PD based on deep learning using midbrain MR images. The system is component of two neural networks. The first one is the Faster R-CNN, which identifies the areas that may be used for PD diagnosis from midbrain MR images. As Parkinsonian images are highly similar, the non-diagnostic areas that can't be used for diagnosis may also be detected by the network during the detection of the areas of interest. Thus, we have three types of identified images, namely, normal, Parkinsonian, and non-diagnostic images. The second neural network is a convolutional neural network defined herein, which classifies the areas identified in the first network. We used MR images of 350 patients for training and testing, and the results showed this system exhibits a higher accuracy in PD diagnosis.

Keywords— Parkinson's Disease; Faster R-CNN; CNN; Deep Learning

1. INTRODUCTION

Parkinson's disease (PD) [1] is a disease related to the nervous system. It was proposed by British doctor James Parkinson in 1817. After whom it was officially named 70 years later. PD typically occurs in older people. Its clinical manifestation includes multiple motor and non-motor disorders. Non-motor disorders are generally difficult to observe, and patients generally appear first on the limbs. The main motor symptoms of PD patients were static tremor, inability to exercise, stiffness, and abnormal gait. Non-motor symptoms include dementia, depression, hallucinations, auditory hallucinations, sleep disorders, olfactory disorders, and autonomic dysfunction symptoms. PD is characterized pathologically by a substantial degeneration and loss of dopaminergic neurons in the substantia nigra, especially in the pars compacta. The causes about pathological changes are still unclear, and there is no cure for PD. In addition, the onset of PD is insidious, and in the early stage, the patient often unaware of their conditions, thus missing the best treatment time. Therefore, reliable diagnosis of PD has become a hot topic of study among researchers.

Received: December 26, 2018

Reviewed: February 26, 2019

Accepted: March 7, 2019

* Corresponding Author

With the advancement of technology, we have been able to classify PD based on molecular biology, genetics, and pathology. Among them, classification according to the type of clinical dyskinesia of patients is commonly used by clinicians. The clinically common Parkinsonian subtypes include akinetic-rigid, tremor-dominant, and the incidence rate of the elderly is high and the pathological development is rapid, the incidence rate of young people is low and the disease progresses slowly. postural instability and gait difficulty, and intermediate (mixed). At present, this classification method is widely used for recognition. It was proposed by Jankovic *et al.*, [2] in 1990. According to the average tremor and average posture gait abnormality ratio in the unified PD rating scale, all patients with PD are classified into posture instability.

Although the main clinical manifestations of PD are severe limb symptoms, researchers generally believe that PD is a neurodegenerative disease that originates in the central rather than the peripheral nervous system. The main pathological change of the disease is progressive degeneration of dopaminergic neurons in the substantia nigra pars compacta and pathological aggregation of Lewy bodies. When the proportion of pathologically deprived neurons is as high as 50–60%, the patient will exhibit clear clinical symptoms. As Parkinson's pathogenesis is complex, it is difficult to diagnose PD.

With the continuous advancement of biomedicine, auxiliary PD diagnostic techniques continue to emerge. For example, computed tomography and magnetic resonance (MR) imaging [3] techniques have been proposed to diagnose PD, among which midbrain MR imaging has proved to be capable of diagnosing PD correctly. Patients usually undergo MR imaging of the midbrain and physicians thereafter examine the images to decide whether the patients have PD. Unfortunately, physicians are required to have significant expertise and medical experience in such human-dependent diagnosis. Furthermore, manual diagnosis is also inefficient.

Fortunately, the great success of deep learning in the field of computer vision has inspired many scholars to apply it to medical image analysis. Professor Wells of Harvard Medical School pointed out in his review that the application of deep learning to solve medical image analysis tasks is a trend in the field. Since 2016, many experts have summarized, reviewed, and discussed the research status and problems of deep learning in medical image analysis. A review published in Medical Image Analysis provides a comprehensive summary of deep learning in medical image classification, detection, segmentation, registration, and retrieval. It can be observed that the application of deep learning in medical imaging is becoming increasingly mature and its application in the classification of ultrasound images and in the diagnosis of lung cancer is a good example.

In our study, an automatic diagnostic algorithm is proposed for PD based on deep learning using midbrain MR images. The algorithm comprises two components. In the first phase, the Faster R-CNN [4] test is used to identify the areas that may be used to diagnose PD. In the second phase, the identified areas are used for PD diagnosis. We applied our algorithm to 70 patients and the results showed that it had an accuracy of 90%.

The structure of the paper is as follows. The second chapter introduces related research and knowledge of convolutional neural networks (CNN) [5], faster R-CNN, and Parkinson's diagnosis. In the third chapter, the proposed algorithm, the processing of the data, and the flow of the algorithm are introduced. In the fourth chapter, the test results of the algorithm are shown. Here, we test the results of the regional detection and the results of the classification. The fifth chapter presents our conclusion.

2. RELATED WORK

To study the Parkinson's diagnostic system, we first review the relevant knowledge on deep learning algorithms. In this section, we introduce CNNs, Faster R-CNN, and Parkinson's diagnosis.

2.1. CONVOLUTIONAL NEURAL NETWORK

Deep learning [6] is a special branch of machine learning, which is an artificial neural network formed by simulating the human brain to analyze problems [7]. Representative of these is the backpropagation(BP) neural network [8]. The BP neural network is a multi-layer network composed of multiple perceptron's. The forward propagation calculates the resulting value and then back propagates the error, training the weight of the network by minimizing the error between the actual value and the expected output value of the network. To process image information, a CNN has been proposed.

CNN is a part of ANN(artificial neural network) that can process images. In 2012, the CNN model AlexNet [9] achieved remarkable results in the ImageNet competition. The model reduced the classification error rate from 26% to 15%. Subsequently, CNN has become an indispensable part of the field of pattern recognition.

The biggest advantage of using CNN is that CNN can automatically extract data features through training data, which avoids the process of manually extracting features from data. Thereby the image can be directly used as an input to the convolutional neural network. In addition, compared with neural networks, convolutional neural networks have the characteristics of local perception and parameter sharing. Due to these two characteristics of convolutional neural networks, the complexity and weight of the network will be reduced, and the convolutional neural network The structure remains invariant for image translational rotation scaling. The traditional CNN mainly includes the following five parts:

Convolutional layer: As the core part of convolution neural network, each layer of convolution consists of a set of convolution kernels. It can learn to change the parameters of the convolution kernel. The main function is to extract image features. Convolution kernels have different sizes. The commonly used convolution kernel sizes are 3×3 and 5×5 . In particular, 3×3 convolution kernels are used more frequently.

Fig. 1 Example of convolution calculation

Activation layer: The main effect of activation layer in CNNs is to enable the network to perform nonlinear modeling. If a CNN contains only the convolutional kernel and fully-connected layer, the network can only express linear mapping, and hence, it is difficult to deal with nonlinear data in a modeling environment. Therefore, For the case of nonlinear data, an activation function can be applied. There are many activation functions that can be applied to different situations, such as Sigmoid [10], Tanh [11], and ReLU [12].

Fig. 2 Sigmoid and ReLU activation functions

Pooling layer: In the CNN network, the pooling layer[13] generally follows the convolutional layer, and integrates the features extracted by CNN in a small neighborhood to reduce the size of the output space. And the pooling layer can also reduce the parameters in the network by reducing the size of the feature map. Usually, the pooling layer obtains the main features of the feature map by down sampling. There are max pooling and average pooling, with the max pooling being more common when processing images.

Fig. 3 Max pooling; the left image is the input feature map and the right image is the max pooling output.

Fully-connected layer: The fully connected layer is similar to the traditional artificial neural network, and its function is to classify the feature extracted by the convolutional layer. The feature map data output from the last layer in the convolutional layer is taken as input, and then back propagated to classify the data. Since the fully connected layer is a neuron that performs connection calculation for each data, generally 80% of the parameters of a network are in the fully connected layer. Therefore, over-fitting is easily generated in the fully connected layer. To this end, in order to reduce the number of fully connected parameters, the researchers proposed an algorithm that replaces the fully connected layer with a 1*1 convolution.

Fig. 4 Fully-connected layer

Dropout: In general, by deepening the network and using more neurons, a CNN can achieve higher performance, but because the training data set is too small to be over-fitted, Hinto *et al.*, [14] proposed the dropout method, whereby, when the model is trained, some nodes do not work, their output is set to 0. Applying the dropout method can effectively prevent overfitting.

Fig. 5: Dropout example; the left image is the standard neural net and the right image is obtained after applying dropout.

2.2. RESNET

Since AlexNet won the championship in the ImageNet competition in 2012, the CNN has become a hot topic of research. Subsequently, large-scale networks with good effects have been proposed every year, such as VGGNet[15], GoogLeNet[16], *etc.*, ReaNet[17] is a classification algorithm that was introduced in 2015. We know that this algorithm is at the top of the ImageNet competition this year. ResNet introduced a residual unit and successfully trained a 152-layer CNN. The parameters of the network are fewer than those of VGG19, but the classification effect is better than that of VGG. It can be observed that the introduction of the residual unit structure can significantly increase training speed of network training and get better network training performance. In the previous network, researchers realized that by deepening of the network structure, the accuracy of the training set decreased. The reason for this phenomenon is that the gradient disappears owing to the increase in the number of network layers. To solve the problem of gradient disappearance, ResNet proposed two kinds of mapping: identity mapping and residual mapping. The residual block of ResNet is shown as follows:

Fig. 6 Residual block of ResNet

From the perspective of information flow, directly inputting the input information to the output protects the integrity of the information. Compared with the middle-layer information transmission in a traditional neural network, ResNet solves the problem of information loss. Owing to its simple structure, ResNet is also used in the fields of detection, segmentation, and recognition.

Fig. 7 Structure of ResNet51

Fig.7 shows the network structure of ResNet51. It can be observed from the Fig that ResNet51 is composed of three layers of residual learning units. Experiments show that after using the residual unit on training set and test set, the performance of the ResNet become better as the number of network layers deepens. thus, we can implement the training and testing phases.

2.3. FASTER R-CNN

Faster R-CNN was proposed by Ross Girshick *et al.*, Which is an object detection algorithm. It first obtains a feature map using the CNN network, and thereafter generates regional recommendations through the region proposal network (RPN), after which the classification and bounding-box regression are performed within the region, and the exact location of the object is generated. Compared with the traditional R-CNN [18] and Fast R-CNN [19], the detection speed of Faster R-CNN is increased by 10 times owing to the use of the RPN module and it is an end-to-end system. The structure is shown in Fig. 8.

Fig. 8 Structure of Faster R-CNN

It can be seen from the above Fig, Faster R-CNN is mainly di-vided into two large modules: The first module is the RPN module for distinguishing objects from the background, and the second module is the detection module. First, we input the image using the CNN to extract features, and thereafter obtain the feature maps corresponding to the image, which is copied into two parts. One part enters the RPN generation object bounding boxes, and the other part enters the region of interest (ROI) pooling layer. The main function in this layer aims to unify the candidate frame feature size of the RPN output and the feature maps of convolutional layer output. Behind ROI layer is a fully-connected layer, which has two functions: classification and regression. When classifying, we use soft-max to classify. In the position regression phase, we use linear regression function to perform regression, and in the case of backpropagation, the two phases of error are back propagated together for optimization.

2.4. PARKINSON'S DIAGNOSIS

PD [20] is a progressive neurodegenerative disease that can destroy brain cells that control exercise. PD currently has no cure. Early diagnosis and treatment are vital in treating PD. The diagnosis of PD proposed in 1991 includes the following criteria:

- (a) May be PD: a person with any of the three symptoms: tremor (stationary or postural), tonic, and slow motion.
- (b) First diagnosed as PD: two of the four main symptoms resting tremor, rigidity, bradykinesia, and posture reflex disorder. Among the three symptoms, namely, asymmetry tremor, tonic, and bradykinesia, one of the three manifestations is evident.
- (c) Diagnosed as PD: Three of the four main symptoms resting tremor, rigidity, bradykinesia, and posture reflex disorder two of which are evident.

The analysis of these diagnostic methods has a reference role in the diagnosis of PD, but there has been no uniform standard. There has been a high misdiagnosis rate in clinical practice. Dr. Stefan Schwarz [21] proposed a method for judging PD using the midbrain substantia nigra. Subsequently, doctors have used this method to diagnose PD. For example, the normal MR images and MR images of patients with PD are shown in Fig. 9.

Fig. 9 Normal and Parkinson MR images; the left images are normal images and the right images are Parkinson images

3. PROPOSED ALGORITHM

As MR images can be used to diagnose PD, we can automate the diagnosis of PD by processing MR images. This paper proposed a diagnostic system for PD based on deep learning using midbrain MR images automatically. The system can diagnose PD based on MR images and it comprises two parts. In the first part, Faster R-CNN is used to detect the diagnostic areas in the MR image, and thereafter classify the discriminable area. In the second part, we define a CNN to classify the diagnostic areas, thus realizing the diagnosis of PD. The flowchart of our method is shown as follows:

Fig. 10 Flowchart of the system

3.1. IDENTIFICATION OF THE DIAGNOSTIC AREAS

When diagnosing PD, physicians must take multiple MR images of patients generally, 30 pictures. However, not every image may be used for diagnosis and we name the area of interest the ‘substantia nigra area,’ which must be identified. Examples of a patient’s MR images are shown as follows:

Fig. 11 MR images of a patient

In Fig. 11, it is evident that each image is different with a different imaging region. Fig. 11(3) and (4) contain the diagnostic area, and we first used the Faster R-CNN algorithm to identify it.

3.2. TRAINING THE FASTER R-CNN NETWORK

3.2.1. CREATING THE TRAINING DATA SET: When training Faster R-CNN network, we need a dataset, we tested 350 patients, with 30 MR images for each patient. We selected five images that showed the diagnostic area most evidently for each patient for annotation. We obtained 1,750 images to be annotated as the training data set. The original image size is 384×384 in our data set. To reduce the amount of calculation, we reduced the image size to 250×250 size based on the center. We manually marked the diagnostic area on the image; the label contains the title of the image, the information of the starting point, and the size of each identified area. We named the diagnostic area SR, and after the comment, we saved all the annotated information in an xml file. Below is an example of an MR image.

MRI image

```
<annotation>
  <folder>MRI_images</folder>
  <filename>Nor_res0.jpg</filename>
  <size>
    <width>250</width>
    <height>250</height>
    <depth>1</depth>
  </size>
  <object>
    <name>SR</name>
    <bndbox>
      <xmin>57</xmin>
      <ymin>102</ymin>
      <xmax>194</xmax>
      <ymax>188</ymax>
    </bndbox>
  </object>
</annotation>
```

Annotation information

Fig. 12 A patient's MR images and annotation information; the red bounding box in the original image represents the position we marked, and in the annotation information, we include the size of the image and the starting and ending positions of the frame. As only one recognition area is required in our image, there is only one object.

After collecting the data, we move on to the next step, training the network.

3.2.2. TRAINING FASTER R-CNN: According to Section 2.3, we know that CNN networks are used to extract image features when we classify images through Faster RCNN. A good feature extraction network can extract robust features, which are very useful in the subsequent classification and regression tasks. Here, we use ResNet51 [22] as the feature extraction network. At the same time, we use the classification model to obtain the weight of the network parameters. To make the training results more accurate and obtain them quickly, we changed the number of anchors generated in the RPN network. As the length-to-height ratio of our image discriminable regions is approximately 2:1, only an anchor of size 2:1 is generated. This reduces the number of anchors and makes detection faster.

As the detected area is slightly different from the manually labeled area, we use the positional correctness (IoU) of the border, which defines the overlap of the two rectangular boxes, to measure the performance of the algorithm.

Fig. 13 IoU Example

In Fig. 13, A indicates the detected area, and B indicates the manually calibrated bounding box. Thus, the overlap degree of A and B, defined as the intersection of A and B, accounts for the ratio of the union of A and B. The formula of IoU is as follows:

$$\text{IoU} = \frac{A \cap B}{A \cup B} \quad (1)$$

It is known that the larger the value of IOU, the better is the detection performance of the algorithm.

3.2.3. TESTING RESULT: After the training, we tested the Faster R-CNN. We divide the data into 80% training images and 20% test images. The average IoU accuracy was 63%. However, as the substantia nigra area contained in the unrecognized picture is similar, when the authenticate R-CNN is used to detect the identified areas, many non-diagnostic areas are detected. Therefore, in the next step of our classification, the non-diagnostic images are classified into a different category. The test step results are as follows:

Fig. 14 Faster R-CNN detection result; Left, input images. Right, the result after Faster R-CNN. (1) the result of normal image detection, (2) the result of Parkinson's image detection, (3) the result of a non-diagnostic image detection.

After the area is tested, the next step is to classify the area. We define a classification convolution network for this purpose.

3.3. CLASSIFICATION OF THE IDENTIFIED AREAS

In the previous section, we used the Faster R-CNN algorithm to identify the diagnostic areas. In this section, we classify all the areas that have been identified. As the substantia nigra area in each MR image is highly similar, three types of identified areas are obtained using the Faster R-CNN algorithm, *i.e.*, normal, Parkinsonian, and non-diagnostic images, which are shown in Fig. 15.

Fig. 15 Normal, Parkinsonian, and non-diagnostic images

As can be seen from Fig. 15, all the MR images are similar; therefore, conventional methods are extremely limited in identifying the characteristics of each image, and conventional machines cannot classify them. Hence, we used a CNN for classification.

As the areas identified using the Faster R-CNN algorithm were not identical in size, we had to resize them. Most of the identified areas were approximately 330×170 in size, and hence, we decided to resize all the images to 330×170 . As our dataset is too small, we expanded the dataset. Dataset expansion methods include rotation, scale transformation, and translation. Here, we select the translation transformation. We capture two images in one image and save them. The method is shown in Fig. 16.

Fig. 16: Image capture

We tested the data from 350 subjects, and the number of identified areas of each type is shown in Table I.

Table I. Number of identified areas of each type

Classification	Amount
Normal	1750
Parkinson	1750
Non-diagnostic	2800

After the data set was established, the classification model will be training using the data set. First, we use Google Net V3 to classify. GoogleNet V3 is a network designed in the ImgeNet competition. It uses the inception structure to improve the classification results. Since GoogleNet V3 is designed on ImageNet, we can classify our data. We changed the output of the network to three categories. We used 80% of the data, and then 20% of the data were used for testing. Since the structure of the googLeNet V3 model is too complicated and there are too many parameters, the overfitting occurs when training our model. So the correct rate of use in the test is only 78.3%. For this reason, we defined the CNN [23,24] model for training using the data. We defined a 10-layer CNN for training, the network structure we defined is shown below:

Fig. 17 Proposed CNN structure

As shown in Fig. 17, We define a ten-layer CNN model with seven convolutional layers in the model. And there are two pooling layers in the middle of the convolution layer for max pooling. In the fully connected layer we define a three-layer fully connected layer for classification. Our network inputs were 330×170 grayscale images. In the first and second convolution layers, we used convolution with 64 3×3 convolution kernels. We set the moving step size of the convolution kernel to 1. When choosing the activation function, we chose ReLU as the activation function in order to reduce the disappearance of the gradient. So we get 64 330×170 feature maps. In the third layer of convolution, to prevent overfitting by reducing the number of parameters, we used 32 3×3 filters to perform convolution, and after the third layer, we set a pooling layer. We use a 2×2 filter for max pooling, and the step size is set to 2. From the fourth to seventh layers of convolution, we used 32 3×3 convolution kernels for convolution, and hence, the output data at the seventh layer were of size $168 \times 85 \times 32$ to reduce the amount of data to prevent overfitting. After the seventh convolutional layer, we set a pooling layer, also using 2×2 filter for max pooling, and using the step size 2 to reduce the input of data to the fully connected layer. After the pooling layer, we expand the data into 1-D data input to the fully connected layer for classification. In the first fully connected layer, we set up 128 neurons to connect and use the ReLU function as the activation function. In the second layer of the fully connected layer, we also used 128 neurons to connect. In the classification layer, since we only have three categories, we set it to three neurons. In network training, we use the back propagation algorithm to train and use cross-entropy loss for optimization.

To test the performance of our model, we implemented our algorithm using the TensorFlow library. And use our Parkinson diagnostic image for testing.

In the training phase. We set the learning rate to 0.0002. In order to fully fit our data, we set the total number of trainings to 20,000 rounds. To prevent over-fitting of the model, we divided 80% of the data into a training set and used the remaining data as a test set to test the model. In addition, we used a dropout of 0.2 to prevent overfitting. The accuracy of the testing is shown in Table II.

Table II. Testing accuracy

Classification	Training data	Testing data	Accuracy
Normal	1400	350	92%
Parkinson	1400	350	89.7%
Non-diagnostic	2240	560	88.2%

From Table II, we can observe that our classification results have an accuracy of approximately 90%. Compared to GoogLeNet V3 model, our result is better.

4. EXPERIMENTAL RESULTS

To test the performance of our algorithm, we implemented the PD diagnosis system. The experimental environment is as follows:

1. Hardware environment: Intel (R) Core i7-6700K CPU at 4.00 GHz
2. System environment: Ubuntu 18.04.01
3. Software environment: Python, TensorFlow [26], Keras[27], python-OpenCV [28], CuDa8.0

This system uses the Keras implementation program to train the Faster R-CNN. Keras is a high-level neural network API. It is based on Tensorflow, Theano, and CNTK libraries and it can easily build deep-learning programs. We used TensorFlow to build a classification program. TensorFlow is a deep-learning framework developed by Google. It supports Linux, Window, and Mac platforms. TensorFlow provides a rich set of deep-learning-related APIs, including basic vector matrix calculations, various optimization algorithms, basic unit implementations for various CNNs, and visual aids.

When testing our algorithm, we used a data set provided by the hospital. There were 70 subjects, including 35 normal people and 35 patients with PD. We separately captured and collected MR images of these subjects. According to the sequence images obtained, we selected 45 images as our samples. After sorting these images, the data were generated for testing. When testing, 27 images of a subject were entered into our network. The position of substantia nigra in the MR image is detected using the Faster R-CNN, and thereafter, the detected substantia nigra regions are classified using our CNN network. The classification of these images results in normal images, Parkinson images, and non-diagnostic images. We calculate the ratio of the number of images in each of these images to the total number to determine whether this person has PD. The MR images used as input are shown in Fig. 18.

Fig. 18 Test images entered as input: 27 MR images per person.

The images in Fig. 18 are MR images captured by a tester. First, with the Faster R-CNN, the diagnosable areas of the images are detected. The test results are shown in Fig. 19.

Fig. 19 Result of the Faster R-CNN test; among these images, it can be judged only from the ninth to thirteenth images whether the patient has PD, and ideally, only the determinable regions of the five images should be detected. However, as the similarity of the images is too high, many areas (18 areas) are detected in the above Fig.

As shown in Fig. 19, we used the Faster R-CNN to detect 18 different areas. Further, we used the CNN classification network to classify the detected areas. The classification results are shown in Fig. 20.

The input images number is 18
The Normal images number is 1
The Parkinson images number is 5
The No-diagnostic images number is 12
The diagnosis result is Parkinson

Fig. 20 CNN classification Result

We input 18 images with detected areas into the CNN network. The CNN classification results indicate that there is one normal image, five Parkinson images, and 12 non-diagnostic images. As the number of Parkinson images is larger than the number of normal images, the result of our diagnosis is Parkinson. We tested 70 testers, and the test results showed that our algorithm has the accuracy of 91.4%.

5. Conclusion

In this paper, we proposed a deep-learning-based algorithm to diagnose PD. By training a CNN, midbrain MR images can be detected and classified to automatically diagnose whether a person has PD. When training the CNN, we chose to use the Faster R-CNN network, because this network ensures faster image detection and higher accuracy. Thus, we can detect the areas of interest in the MR images. Simultaneously, we define a network that can implement image classification to classify images with different characteristics. These are the two essential aspects of diagnosing PD.

In the study of PD, we learned that the pathological features of the disease are characterized by significant degeneration of dopaminergic neurons in the substantia nigra, while a large number of Lewy bodies are precipitated in the remaining neurons, leading to dopaminergic projection of collaterals. This lesion can be observed with an MR image, and the MR image captures the human midbrain with its precise positioning and non-invasive advantage. Owing to its advantages, MR image analysis is one of the most commonly used methods to diagnose PD. To mitigate the influence of the difference in doctors' professional knowledge and actual experience on the judgment results, we proposed an automatic diagnostic system for PD based on deep learning using midbrain MR images.

The first step of our algorithm is to detect the position of the substantia nigra dopaminergic neurons in the image of the human midbrain captured using MR imaging. We can only analyze and judge the lesions if we determine the location of the lesion. In this step, we use the Faster R-CNN algorithm to detect the image. Although there are some subtle differences in the brain structure of each person, the MR images captured by different people are different, and the images of different degrees of lesions are also different, but the Faster R-CNN can accurately and quickly find the black matter in these different images. The location is detected accurately. The second step is to classify the detected locations. Here, we define our own CNN structure as per our requirements. According to the difference in substantia nigra in MR images, we can determine whether a person has PD. However, it is impossible to diagnose PD owing to the different definitions of the image. Therefore, through our own network training, these images can be classified into normal, Parkinson, and non-diagnostic images. Thus, the patient's condition can be automatically determined using deep learning. The experimental results show that the use of this method effectively improves the diagnostic accuracy of PD, which has a good auxiliary effect on the doctor's judgment of the condition.

ACKNOWLEDGEMENT

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2018-2017-0-

01630) supervised by the IITP(Institute for Information & communications Technology Promotion)

REFERENCES

- [1] Hughes, Andrew J., et al., "Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases." *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 55, no. 3, (1992), pp.181-184, available online: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1014720/>
- [2] Jankovic, Joseph, et al., "Variable expression of Parkinson's disease A base- line analysis of the DAT ATOP cohort." *Neurology*, vol. 40, no. 10, (1990), pp: 1529-1529, <https://www.ncbi.nlm.nih.gov/pubmed/2215943>.
- [3] Quattrone, Aldo, et al., "MR imaging index for differentiation of progressive supranuclear palsy from Parkinson disease and the Parkinson variant of multiple system atrophy.", *Radiology*, vol. 246, no. 1 (2008), pp:214-221, available online: <https://www.ncbi.nlm.nih.gov/pubmed/17991785>.
- [4] Ren, S., He, K., Girshick, R., & Sun, J., "Faster r-cnn: Towards real-time object detection with region proposal networks." In *Advances in neural information processing systems*, (2015), pp: 91-99, available online: <http://papers.nips.cc/paper/5638-faster-r-cnn-towards-real-time-object-detection-with-region-proposal-networks.pdf>.
- [5] LeCun, Yann, and Yoshua Bengio, "Convolutional networks for images, speech, and time series.", *The handbook of brain theory and neural networks*, vol. 3361, no. 10, (1995), https://www.researchgate.net/profile/Yann_Lecun/publication/2453996_Convolutional_Networks_for_I_mages_Speech_and_Time-Series/links/0deec519dfa2325502000000.pdf.
- [6] Praveen Ch, ra, "A Survey on Deep Learning its Architecture and Various Applications.", *Asia-Pacific Journal of Neural Networks and Its Applications*, vol. 1, no. 2, (2017) September.
- [7] Abraham, Ajith. "Artificial neural networks." *handbook of measuring system design*, (2005), <https://onlinelibrary.wiley.com/doi/pdf/10.1002/0471497398.mm421>.
- [8] Hecht-Nielsen, Robert, "Theory of the backpropagation neural network.", *Neural networks for perception*, (1992), pp:65-93, <https://doi.org/10.1016/B978-0-12-741252-8.50010-8>.
- [9] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton, "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems*, (2012), pp:1097-1105, available online: <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>.
- [10] Han, Jun, and Claudio Moraga, "The influence of the sigmoid function parameters on the speed of backpropagation learning." *International Workshop on Artificial Neural Networks*. Springer, Berlin, Heidelberg, (1995), pp:195-201
- [11] Malfliet, Willy, and Willy Hereman. "The tanh method: I. Exact solutions of nonlinear evolution and wave equations." *Physica Scripta*, Vol.54, No.6, (1996), <http://iopscience.iop.org/article/10.1088/0031-8949/54/6/003/meta>
- [12] Li, Yuanzhi, and Yang Yuan, "Convergence analysis of two-layer neural networks with relu activation", *Advances in Neural Information Processing Systems*, 2017, pp:597-607
- [13] Scherer, Dominik, Andreas Müller, and Sven Behnke, "Evaluation of pooling operations in convolutional architectures for object recognition", *Artificial Neural Networks-ICANN*, Springer, Berlin, Heidelberg, (2010), pp:92-101, https://link.springer.com/chapter/10.1007%2F978-3-642-15825-4_10
- [14] Srivastava, Nitish, et al., "Dropout: a simple way to prevent neural networks from overfitting", *The Journal of Machine Learning Research*, Vol.15, No.1, (2014), pp: 1929-1958, <https://www.cs.toronto.edu/~hinton/absps/JMLRdropout.pdf>
- [15] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition", *arXiv preprint arXiv*, (2014), pp:1409-1556, available online: <https://arxiv.org/abs/1409.1556v6>
- [16] Szegedy, Christian, et al. "Going deeper with convolutions", *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2015), pp:1-9, <https://ieeexplore.ieee.org/document/7298594/>
- [17] He, Kaiming, et al., "Deep residual learning for image recognition", *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2016), pp:770-778, <https://arxiv.org/abs/1512.03385>
- [18] Girshick, Ross, et al. "Rich feature hierarchies for accurate object detection and semantic segmentation", *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2014), pp:580-587, <https://arxiv.org/pdf/1311.2524.pdf>
- [19] Girshick, Ross. "Fast r-cnn", *Proceedings of the IEEE international conference on computer vision*, (2015), pp:1440-1448, https://www.cv-foundation.org/openaccess/content_iccv_2015/papers/Girshick_Fast_R-CNN_ICCV_2015_paper.pdf
- [20] Gelb, Douglas J., Eugene Oliver, and Sid Gilman, "Diagnostic criteria for Parkinson disease", *Archives of neurology*, Vol.56, No.1, (1999), pp:33-39, <https://jamanetwork.com/journals/jamaneurology/article-abstract/774675>

- [21] Schwarz, Stefan T., et al. "The 'Swallow Tail' Appearance of the Healthy Nigrosome—A New Accurate Test of Parkinson's Disease: A Case-Control and Retrospective Cross-Sectional MRI Study at 3T", *PloS one*, Vol.9, No.4, (2014): e93814, <https://www.ncbi.nlm.nih.gov/pubmed/24710392>
- [22] Tang, Wenyi, Bin Liu, and Nenghai Yu, "Deep Scale Feature for Visual Tracking", *International Conference on Image and Graphics*, Springer, Cham, (2017), https://link.springer.com/chapter/10.1007/978-3-319-71607-7_27
- [23] Hu, Wei, et al., "Deep convolutional neural networks for hyperspectral image classification", *Journal of Sensors* 2015, (2015), <http://dx.doi.org/10.1155/2015/258619>
- [24] Cireşan, Dan, Ueli Meier, and Jürgen Schmidhuber, "Multi-column deep neural networks for image classification", *arXiv preprint arXiv:1202.2745* (2012), <https://arxiv.org/abs/1202.2745>
- [25] Jun Mo Jeong, Se Jin Choi, Chiyong Kim3, "Correcting Method of Numerical errors of Soft-max and Cross-entropy according to CNN's Output value".*Asia-Pacific Journal of Neural Networks and Its Applications*.Vol.1, No.1, May(2017)
- [26] Abadi, Martín, et al., "Tensorflow: a system for large-scale machine learning", *OSDI*. Vol. 16, (2016), <https://arxiv.org/abs/1605.08695>
- [27] Chollet, François. "Keras." (2015), <https://keras.io/>
- [28] https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_tutorials.html