Proceedings of the 24th International Conference on Automation & Computing, Newcastle University, Newcastle upon Tyne, UK, 6-7 September 2018

Detection of Parkinson Disease in Brain MRI using Convolutional Neural Network

Pir Masoom Shah[1], Adnan zeb[2], Uferah Shafi[3], Syed Farhan Alam Zaidi[4], Munam Ali shah[5]
Department of Computer Science
COMSATS University Islambad, Islamabad Pakistan[1245]
NUST University Islamabad, Pakistan[3]
pirmasoomshah@gmail.com[1], adnanzeb933@gmail.com[2], uferahshafi@gmail.com[3]
syedfarhanalam1993@gmail.com[4],mshah@comsats.edu.pk[5]

***Abstract*—Parkinson Disease (PD) is one of the most critical progressive neurological diseases which mainly affects the motor system. The accurate diagnosis of PD has been a challenge to date, mainly due to the close relevance of PD to other neurological diseases. These close characteristics are the reasons that cause 25% inaccurate manual diagnosis of PD. In this paper, we present a Convolutional Neural Network (CNN) based automatic diagnosis system which accurately classifies PD and healthy control (HC). Parkinson's Progression Markers Initiative (PPMI) provides publically available benchmark T2-weighted Magnetic Resonance Imaging (MRI) for both PD and HC. The mid-brain slices of 500, T2-weighted MRI are selected and aligned using image registration technique. The performance of the proposed technique is evaluated using accuracy, sensitivity, specificity and AUC (Area Under Curve). The detailed comparison in the result section shows that the CNN archived a better performance from 3% - 9% in terms of accuracy, sensitivity, specificity, and AUC when compared to the some existing techniques.**

Keywords—Parkinson Disease, MRI, Deep Learning, Convolutional Neural Network, CNN

I. INTRODUCTION

Parkinson disease (PD) is a non-curable progressive neurological disorder, which affects the motor system of the human brain. Early medication can give temporary relief to the patients and slows down PD progression [1]. It occurs due to the neurological disorder [2]. The substantia nigra is a thelmic region located in the midbrain, which contains a high level of dopamines neurons. Dopamines are the specific type of chemicals released by neurons for sending signals to other adjacent neurons in the brain. PD occurs when the degeneration of these dopamine neurons starts in substantia nigra [3] [4]. This results in resting tremor, bradykinesia and rigidity problems in patients [5]. Some other symptoms include fatigue, anxiety, depression, slowness in thinking and voice disorder [6] [42].

After Alzheimer, PD is the second largest neurological disorder which is common [7] in aged people. The main cause of PD is still unknown but involves genetic and environmental factors [8]. Due to the lack of medical labs, it is mostly diagnosed at advanced stages [9]. In order to diagnose, experts use history and neurological investigations [10]. However, this approach is less accurate as similar symptoms of other neurodegeneration diseases exist. It is mostly diagnosed when dopamine chemicals are lost heavily.

It is believed that 25% of diagnoses are incorrect [11]. The accurate detection of PD is still a challenging task. If a patient has PD but inaccurately diagnosed as healthy, the disease may progress and become difficult to control. It can be diagnosd by several clinical tests. But as it is linked with biological changes in the brain, the detection with the visual image analysis is an appropriate method. Positron Emission Tomography (PET) and Single Photon Emission Computed Tomography (SRECT) are two imaging techniques, commonly used in PD diagnosis. Also, few studies demonstrate that these two imaging strategies are capable of PD diagnosis [12]. However, due to their invasiveness [13] and high cost, doctors don't prefer using these two techniques. Magnetic Resonance Imaging (MRI) is a non-invasive imaging technology, which is rarely used in PD detection. However, the recent advancements in MRI made the detection comparatively easier.

Some machine learning techniques have also been used for the detection of PD in MRI images including Bayes [14], Decision Tree [15], Support Vector Machine (SVM) [14, 16, 17, 18, 19], and Artificial Neural Network (ANN) [12]. These machine learning techniques are mainly based on hand-crafted features, where most relevant features are selected manually. Less relevant features are normally eliminated using different dimensionality reduction techniques. Convolutional Neural Network (CNN) is a Deep Neural Network (DNN) and biologically inspired technique, that doesn't require hand-crafted features. Self-feature learning capability of CNN enables it to outperform many state-of-the-art results in computer vision tasks. For example, AlexNet [20], ZFNET [21], VGGNet [22], and RESNet [23] on ImageNet [24] dataset, outperformed other techniques in image classification challenge.

Many studies attempted CNN based model for biomedical image classification and achieved state-of-the-art results. Recently, Billones, Ciprian D, et al. [39] fine-tuned VGGNet for Alzheimer detection and achieved 91.85% accuracy. Similarly, Dou. Qi, et al. [40] achieved 93.16% accuracy on Cerebral Micro-bleeds in MR images. Due to the proven performance of CCN with MR images, we used CNN as a classifier for PD detection and achieved promising state-of-the-art results. Fig.1

shows the overall operation of the proposed system. Initially, MRI data of 500 subjects are obtained and specific slices are made aligned with image registration technique. Furhthermore, mid-brain are cropped from each selected slice using the Free Hand Region Of Interest (ROI) technique. The resultant images are divided into training, testing and validation sets. The CNN is trained, validated and tested on training, validation and testing sets, respectively. Fig. 1 shows the system diagram of the proposed model.

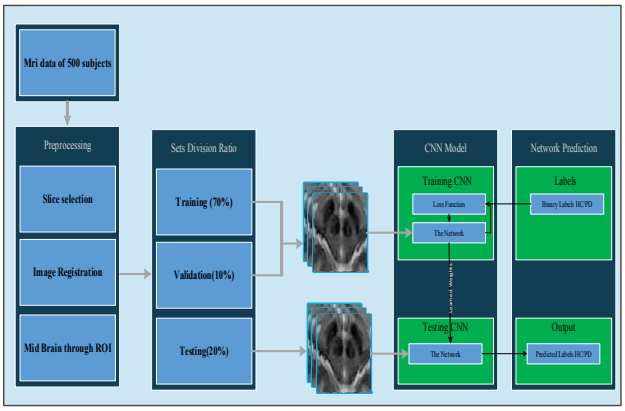

Figure 1. System Diagram

In this study, our contribution includes as per the following:

- A novel 2D- CNN based approach is used to detect Parkinson disease in MRI.
- While dealing with limited data, we use Drop Out layers that reduce over-fitting.
- For decreasing computational cost, we extract specific patches from MRI.

II. Related Work

In this section, we review some existing machine learning techniques for diagnoses of Parkinson Disease.

Olanrewaju et al. in the article [25], proposed machine learning based technique for diagnosis of PD and developed Multilayer Feed Forward Neural Network (MLFNN). They used a data set, available in Oxford Parkinson disease datasets. The dataset consists of voice measurements of 31 people with 23 patients. They used 8 attributes which are based on frequency (tremor). They used a total of 8 input and 10 hidden nodes. For classification, they used the k-mean algorithm. The simulation result showed that they achieved a sensitivity of 83.3%, specificity of 63.6% and accuracy up to 80%. However, this method is not validated on real data.

Prashanth et al. in the article [26], used promoter phase which is an olfactory loss and sleeping disorder feature in their research. They used 40 items from the University of Pennsylvania Smell Identification Test (UPSIT) and sleep behavior from Rapid Eye Questioner (RBDSQ) from PPMI datasets. SVM and classification tree were used for PD detection. Both these techniques achieved 84.26% accuracy. However, dopaminergic imaging features were not used optimally and noiseless data improved performance and accuracy of the classifier.

Ghanad and Ahmadi in an article [27], proposed another model for Parkinson detection. This model used Particle Swarm Optimization (PSO) in order to extract optimal features from data while Naive Bayes for classification. They obtained the dataset from the UCI data repository. In this study, 197 items and 23 attributes were used in the data set. Simulation result showed 97.5% accuracy rate. However, this model performed poorly on large datasets.

Prashanth et al. in the article [28], worked on non-motor features for PD diagnosis. The non-motor features consist of Rapid Eye Movement (REM), sleep behavior disorder, and olfactory loss. In this study, they used non-motor features in combination with cerebrospinal fluid measurement and dopaminergic imaging markers features. The dataset was obtained from (PPMI) database. This dataset consists of 183 normal and 401 patients of Parkinson disease. They used Naive Bayes, SVM, Boosted Tree and Random Forest for classification. The simulation result showed 96.4% accuracy rate of SVM. It was found that the combination of different non-motor features yielded better results. However, data contained imbalanced class distribution.

Avci et al. in the article [29], proposed Genetic Algorithm (GA), Walvet Kernet (WK), and Extreme Learning Machine (ELM) for automated diagnosis of Parkinson disease. In this model, Single Layer Neural Network (SLNN) was used and trained with ELM. The dataset was obtained from the UCI data repository. The dataset consists of voice measurement of 31 people including 23 patients and a total of 192 voices. The proposed WK-ELM model used three adjustable parameters. The optimal values for these parameters were then calculated through GA. The simulation results showed an accuracy rate of 96.81%. However, data contained imbalanced class distribution.

In the article [30] Joshi et al. worked on Alzheimer Disease (AD) and Parkinson. In this work, an automated diagnosis system based on machine learning and the neural network was proposed for Alzheimer and Parkinson detection. This work mainly concerned with the study of the most influencing factors that cause Alzheimer and PD. The chi-square evaluation, one R attribute evaluation, and symmetrical uncertain attribute evaluation schemes were used to identify the most influencing factors. The dataset was obtained from Alzheimer Disease Research Center (ADRC) which contains data of 890 patients, in which 65% for Alzheimer and 40% for PD. For classification, 6 different techniques such as Decision Tree, Bagging, Boosted Tree, Random Forest, RBF Neural Network, Multi-layer Perceptron and Neural Network were used. This work found that the most influencing factors involved in AD and PD are age, genes, and alcohol. The result showed 99.25% accuracy for Random Forest and Multi-layer Perceptron. However, these approaches were not applied to different datasets.

Pereira et al. in an article [31], performed a handwritten examination of 55 people where 37 were PD patients. They developed their own dataset from this handwritten examination that contained images of various characters. The data set was

collected at Faculty of Medicine, Botucatu University, Brazil. The author proposed men relative tremor technique which calculates the amount of tremor from given data. The Naive Bayes classifier was used which achieved 78.9 % accuracy. However, this approach didn't use meander images.

Al-fatlawi et al. in the article [32], proposed Deep Belief Network (DBN) for diagnosis of PD. In this work, a data set PDD was obtained from the UCI data repository that contains 195 voice recordings of 31 people and 16 attributes. The proposed DBN yielded 94% accuracy. However, this work didn't calculate harming probability.

C. Salvatorea et al. in [16] used MRI scans of 84 patients. The dataset contains three parts, PSP (Progressive Supranuclear Palsy), PD and Healthy Control (HC). PSP is a disease similar to Parkinson But is less reactive to medication and its progression speed is higher than the PD. The author used Principal Component Analysis (PCA) as feature extractor and SVM as classification technique. The achieved classification accuracy, specificity and sensitivity of PD vs HC are 85.8%, 86% and 86%, PSP vs HC are 89.1%, 89.1%, and 89.5% while PSP vs PD are 88.9%, 88.5%, and 89.5% respectively. However, this technique was not tested on a benchmark dataset.

S Haller et al. in [17] used Susceptibility-Weighted Imaging (SWI) scans of 36 patients. SWI is a new technique in MR imaging which shows the susceptible variation of di erent tissues such as iron, blood, and calcification with the help of contrast enhancement. In this study, the authors proposed SVM for classification of Parkinson and Parkinsonisms at the individual level. They used their own data set which contained a total of 36 people, of which 16 were PD patients while 20 had the other forms of Parkinson. They achieved 86% of classification accuracy. However, they used a small dataset.

Dan Long et al. in [13] used SVM to classify early PD and HC. In their work, they used multimodal MR imaging for each PD and HC subject. They collected functional magnetic resonance imaging (rsfMRI) and structural images both against each subject. From rsfMRI, they extracted the features at three di erent stages RFCS(regional functional connectivity strength), ReHo (regional homogeneity) and ALFF (amplitude of low-frequency fluctuations). Moreover, in structural images, they extracted volume of White Matter (WM) and Gray Matter (GM). Leave-One-Out Cross-Validation (LOOCV) was used to evaluate the performance of SVM. Their model obtained 86.96% accuracy, 92.59% of specificity and 78.95% of sensitivity. They used their own data set that contained 46 subjects where 19 were PD while 27 were HC.

In an article [12], Paul D. proposed ANN based approach for diagnosis of PD. He worked on SPECT and TRODAT imaging. The dataset was comprised of a total of 175 subjects in which 81 were PD while 94 were HC. The images were further processed for extracting the striatal and striatum pixel values. These extracted pixel values were given to ANN as input. The author obtained a total of 94.4% accuracy, 97.5% specificity and 91.4% sensitivity. The author in [41] used SVM on MRI data as a classifier and randomized logistic regression with LOOCV as a feature extractor. The author achieved the mean accuracy up to 80%.

III. Methods

A. Data Acquisition

The dataset used in this research is obtained from PPMI. PMMI is a landmark with the aims to diagnose Parkinson at its earlier and find valid biomarkers. PPMI is an initiative, which provides huge data containing the leading collection of clinical, imaging and biologic samples. The samples are publicly available on (http://www.ppmiinfo.org). A total of 500, T2 weighted MRI scans were downloaded from PPMI in Digital Imaging and Communications in Medicine (DICOM) format. The data contains 250 MRI scans of PD while 250 for HC. Furthermore, data is divided into training, validation and testing sets with a ratio of 70%, 10%, and 20% respectively.

B. Preprocessing

All the MRI scans are downloaded in DICOM format. We convert images from DICOM to JPEG format by using software package of DICOM-to-JPEG, which is publicly available on www.dicomapps.com/dicomtojpeg/index.html. Moreover, slice no. 22 is collected against each patient's data as in our case, only slice no. 22 contains the accurate image of substantial nigra which is associated with PD.

Intensity-based image registration is performed by using MATLAB Image Registration Suit, on the sack of slice No. 22 against each patient for alignment of images according to a base image. To remove the unwanted information from the images which may mislead our network to learn unnecessary features, we cropped the midbrain by 100x100 window to obtain the perfect image of substantial nigra and were finally input to CNN.

IV. The convolutional neural network

CNN is mainly a data-driven approach which is specifically designed to work with two-dimensional data. CNN learns hidden features, which is difficult to express otherwise. CNN based approaches are biologically inspired techniques and have many advantages over other techniques. It helps us to visualize these spatial learned features. Due to its spatial nature, it dramatically reduces the number of hyper-parameters that need to be learned.

A. Proposed Network Architecture

The proposed system receives MR images as input, which is eventually labeled as PD or HC. The model contains a total number of 8 major layers. The order of these layers are as follows, convolution 1, max-pool 1, convolution 2, max-pool 2, convolution 3, dense layer 1, dense layer 2, and an output layer. The kernel size of all convolutional layers (Convt 1, Convt 2, Convt 3) and Max-pooling layers (Max pool 1, Max pool 2) are 3x3 which generate 32 feature maps. The CNN learns these feature maps, which enable CNN to discriminate between PD and HC in MR images.

Usually, in CNN the convolutional layers are followed by pooling layers. In the proposed model, two max-pooling layers of Max-pool 1 and Max-pool 2 come after Convt 1 and Convt 2 convolutional layers with stride 1 and padding 0. Since the

proposed model consists of three dense layers of Dense 1, Dense 2 and Dense 3(Output layer), the last convolutional layer (Convt 3) is connected to Dense 1 which is composed of 30 neurons and is further connected to 10 neurons of the Dense 2 layer. Eventually, the Dense 2 layer is fully connected to probabilistic (softmax) output Dense 3 layer with 2 neurons.

The output of the presented network depicts a probability of membership of each image to the corresponding PD or HC classes. Table.3 shows the total number of layers, their output, and the parameter used in the concerned layer. The parameters shown in table.3 are calculated as follows. Let suppose the first convolutional layer, the input plane for the first layer is 1, and the output planes are 32. The kernel size is 3x3 and 32 are the biases, at the 1st convolutional layer, the total parameters on this layers are 1*32*3*3+32 = 320. On the second convolutional layer, the input and output planes for this layer are 32, while the kernel size is 3x3, and the biases are 32 that makes 32*32*3*3+32 = 9,248 parameters. Similarly, for the third convolutional layer 32*32*3*3+32 = 9,248 parameters. For the 1st dense layer, the parameters are calculated as, the total input plans for the layers are 14,112 while the output planes are 30 with a total number of 30 biases that makes 1,4112*30 + 30 = 4,233,90 parameters, and thus same is the procedure for 2nd and 3rd dense layers.

V. Results And Anylasis

In this section, we discuss the performance of the proposed network on Parkinson dataset.

A. Performance Measures

We use accuracy, sensitivity, specificity and AUC to test the performance of the proposed network. The Accuracy, sensitivity and specificity are calculated as following in equation 1, 2 and 3, respectively.

$$ACC = \frac{TP + TN}{TP + FP + FN + TN} \times 100\% \quad (1)$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100\% \quad (2)$$

$$Specificity = \frac{TN}{FP + TN} \times 100\% \quad (3)$$

According to confusion matrix True Positive (TP) rate means that the subjects with PD are accuratly classified as PD. Similarly, the True Negative (TN) rate means that the healthy subjects are accuratly classified as healthy. False Positive (FP) rate means that the subjects belong to healthy classes but the classifier misclassify as PD. FN are the numbers of false negative which means that the samples belong to PD class but the classifier misclassified as healthy. The Receiver-Operating Characteristic (ROC) is the graphical representation of a logical model. The curve of ROC shows TP and FP rate. AUC is the area under the ROC curve, which is one of the best methods for evaluation of a model.

TABLE.1. COMPARISON TABLE

Technique	Dataset type	Accuracy
SVM [37]	MRI	92.35%
GA-ELM [3]	MRI	89.22%
SVM [38]	MRI	86.67 %
SVM [17]	SWI MRI	86 %
SVM [13]	MRI Multimodal scans	86.96 %
SVM [16]	MRI Scans	89 %
RVM [36]	PET scans	90%
Decision tree [15]	3-T MR imaging	92%
Nave Bayes [14]	MRI	93 %
ANN [12]	SPECT and TRODAT imaging	94 %
CNN (Proposed)	**MRI Scans**	**96%**

TABLE 2. CLASSIFICATION PERFORMANCE

Classification performance			
Accuracy	**Specificity**	**Sensitivity**	**AUC**
96±2	96.87 ±3.13	95.83±0	99.5±0.5

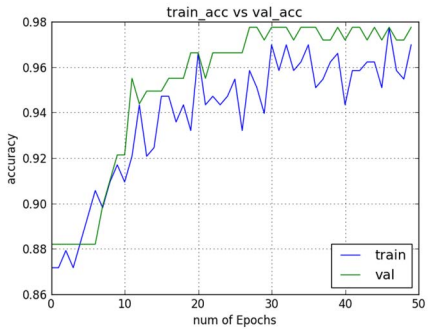

Figure.2. Training vs validation accuracy

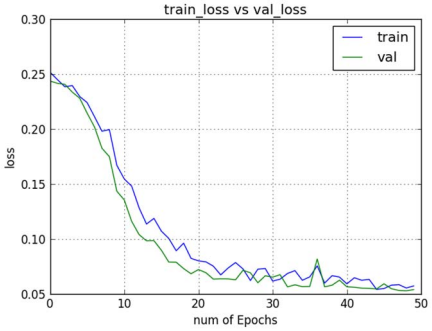

Figure .3. Training vs validation loss

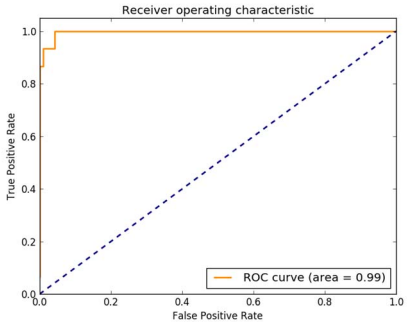

Figure.4. ROC

TABLE 3. NETWORK PARAMETERS

Layer Name	Output shape	Parameters
Convolution2D	(None, 32, 98, 98)	320
Activation	(None, 32, 98, 98)	0
MaxPooling2D	(None, 32, 49, 49)	0
Activation	(None, 32, 49, 49)	0
Convolution2D	(None, 32, 47, 47)	9248
Activation	(None, 32, 47, 47)	0
MaxPooling2D	(None, 32, 23, 23)	0
Activation	(None, 32, 23, 23)	0
Dropout	(None, 32, 23, 23)	0
Convolution2D	(None, 32, 21, 21)	9248
Flatten	(None, 14112)	0
Dense	(None, 30)	423390
Activation	(None, 30)	0
Dropout	(None, 30)	0
Dense	(None, 10)	310
Activation	(None, 10)	0
Dropout	(None, 10)	0
Dense	(None, 2)	22
Activation	(None, 2)	0
Total parameters: 442538		

B. Experimental Setup

CNN is implemented on NVIDIA GeForce 940MX CUDA enabled GPU using keras. Keras is a high level Deep Neural Network (DNN) API written in python which has the capability of running Theano and Tensorflow. Theano and Tensorflow are both DNN libraries. In this study, we use CNN as a sequential model with Theano library.

C. Results and Discussion

Several experiments are performed with different network settings. Network settings include kernel size, batch size, stride and padding. During training of the model, validation and training accuracies are recorded after each epoch. Each time after training the model, is tested on the test set. In all experiments, the classification accuracy fluctuates in between 95 to 98%. Fig.2 shows training vs validation accuracy where x-axis represents the number of epochs while y-axis represents accuracy. The green line shows the validation accuracy while blue line represents the training accuracy. Fig.3 demonstrates the training vs validation loss where x-axis shows the number of epochs and the y-axis shows the loss. Fig.4 shows the ROC and AUC.

D. Performance evaluation

As we mentioned earlier, several experiments are performed with different network settings. Network settings include number of layers, size of input, and other network parameters. In all experiments, the classification accuracy fluctuates in between 94 to 98%. Table.2 shows the mean classification accuracy, sensitivity, specificity and AUC.

The proposed model shows better accuracy when compared to SVM, RVM, Decision tree, ANN,GA-ELM and other machine learning techniques on same dataset. Table .1 represents the comparison of the proposed model with other techniques. Generally, these techniques work on hand-crafted features. Mostly, these machine learning techniques have the risk of missing the most relevant features to the problem. Less appropriate features are normally eliminated by using di□erent dimensionality reduction techniques. Due to the complex manifestation of PD, it is essential to select the most appropriate features related to the disease. The CNN does not requires the handcrafted features. The self-learning capability of CNN enables it to outperform other existing techniques for the same problem and produces higher accuracy rate.

VI. Conclusion

In this study, we proposed a customized CAD based CNN architecture to classify MRI patches of Parkinson and healthy patterns. The proposed network with 3 convolutional layers learns the patterns from training samples of benchmark PPMI dataset in an e□cient way which subsequently improved the accuracy. The results demonstrate that our network is capable of learning accurate features of Parkinson disease automatically. During the experimentation, we found that the limited dataset was a major issue, leading the CNN model towards overfitting. However, with proper design and use of dropout layers in the network, we avoided the overfitting problem.

References

[1] Shetty, Sachin, and Y. S. Rao. ”SVM based machine learning approach to identify Parkinson’s disease using gait analysis.” Inventive Computation Technologies (ICICT), International Conference on. Vol. 2. IEEE, 2016.

[2] Michel, Patrick P., Etienne C. Hirsch, and Stphane Hunot. ”Understanding dopaminergic cell death pathways in Parkinson disease.” Neuron 90.4 (2016): 675-691.

[3] Pahuja, Gunjan, and T. N. Nagabhushan. ”A novel GA-ELM approach for Parkinson’s disease detection using brain structural T1-weighted MRI data.” Cognitive Computing and Information Processing (CCIP), 2016 Second International Conference on. IEEE, 2016.

[4] Chinta, Shankar J., and Julie K. Andersen. ”Dopaminergic neurons.” The international journal of biochemistry & cell biology 37.5 (2005): 942-946

[5] Little, Simon, et al. ”Adaptive deep brain stimulation in advanced Parkinson disease.” Annals of neurology 74.3 (2013): 449-457.

[6] Shulman, L. M., et al. ”Non-recognition of depression and other nonmotor symptoms in Parkinson’s disease.” Parkinsonism & related disorders 8.3 (2002): 193-197.

[7] Obeso, Jose A., et al. ”Missing pieces in the Parkinson’s disease puzzle.” Nature medicine 16.6 (2010): 653-661.

[8] Pringsheim, Tamara, et al. ”The prevalence of Parkinson’s disease: A systematic review and metaanalysis.” Movement disorders 29.13 (2014): 1583-1590.

[9] Fahn, Stanley. ”Description of Parkinson’s disease as a clinical syndrome.” Annals of the New York Academy of Sciences 991.1 (2003): 1-14.

[10] Jankovic, Joseph. ”Parkinsons disease: clinical features and diagnosis.” Journal of Neurology, Neurosurgery & Psychiatry 79.4 (2008): 368-376.

[11] Tolosa, Eduardo, Gregor Wenning, and Werner Poewe. ”The diagnosis of Parkinson’s disease.” The Lancet Neurology 5.1 (2006): 75-86.

[12] Acton, Paul D., and Andrew Newberg. ”Artificial neural network classifier for the diagnosis of Parkinson’s disease using [99mTc] TRODAT-1 and SPECT.” Physics in medicine and biology 51.12 (2006): 3057.

[13] Long, Dan, et al. ”Automatic classification of early Parkinson’s disease with multi-modal MR imaging.” PloS one 7.11 (2012): e47714.

[14] Morales, Dinora A., et al. ”Predicting dementia development in Parkinson’s disease using Bayesian network classifiers.” Psychiatry Research: NeuroImaging 213.2 (2013): 92-98.

[15] Nair, Shalini Rajandran, et al. ”A decision tree for differentiating multiple system atrophy from Parkinsons disease using 3-T MR imaging.” European radiology 23.6 (2013): 1459-1466.

[16] Salvatore, Christian, et al. ”Machine learning on brain MRI data for differential diagnosis of Parkinson’s disease and Progressive Supranuclear Palsy.” Journal of neuroscience methods 222 (2014): 230-237.

[17] Haller, Sven, et al. ”Differentiation between Parkinson disease and other forms of Parkinsonism using support vector machine analysis of susceptibility-weighted imaging (SWI): initial results.” European radiology 23.1 (2013): 12-19.

[18] Duchesne, Simon, Yan Rolland, and Marc Vrin. ”Automated computer differential classification in Parkinsonian syndromes via pattern analysis on MRI.” Academic radiology 16.1 (2009): 61-70.

[19] Focke, Niels K., et al. ”Individual voxelbased subtype prediction can differentiate progressive supranuclear palsy from idiopathic Parkinson syndrome and healthy controls.” Human brain mapping 32.11 (2011): 1905-1915

[20] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. ”Imagenet classification with deep convolutional neural networks.” Advances in neural information processing systems. 2012.

[21] Zeiler, Matthew D., and Rob Fergus. ”Visualizing and understanding convolutional networks.” European conference on computer vision. Springer International Publishing, 2014.

[22] Simonyan, Karen, and Andrew Zisserman. ”Very deep convolutional networks for large-scale image recognition.” arXiv preprint arXiv:1409.1556 (2014).

[23] He, Kaiming, et al. ”Deep residual learning for image recognition.” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.

[24] Russakovsky, Olga, et al. ”Imagenet large scale visual recognition challenge.” International Journal of Computer Vision 115.3 (2015): 211252.

[25] R. F. Olanrewaju, N. S. Sahari, A. A. Musa, and N. Hakiem, Application of neural networks in early detection and diagnosis of Parkinsons disease, 2014 Int. Conf. Cyber IT Serv. Manag., pp. 7882, 2014

[26] R. Prashanth, S. Dutta Roy, P. K. Mandal, and S. Ghosh, Parkinson s disease detection using olfactory loss and REM sleep disorder features, Annu. Int. Conf. IEEE, pp. 57645767, 2014

[27] N.K.GhanadandS.Ahmadi,CombinationofPSOAlgorithmandNaive Bayesian Classification for Parkinson Disease Diagnosis, Adv. Comput. Sci. an Int. J., vol. 4, no. 4, pp. 119125, 2015.

[28] R. Prashanth, S. Dutta Roy, P. K. Mandal, and S. Ghosh, High-Accuracy Detection of Early Parkinsons Disease through Multimodal Features and Machine Learning, Int. J. Med. Inform., vol. 90, pp. 1321, 2016.

[29] D. Avci and A. Dogantekin, An Expert Diagnosis System for Parkinson Disease Based on Genetic Algorithm-Wavelet Kernel-Extreme Learning Machine, Parkinsons. Dis., vol. 2016, 2016.

[30] S. Joshi, D. Shenoy, V. S. G.G., P. L. Rrashmi, K. R. Venugopal, and L. M. Patnaik, Classification of Alzheimers Disease and Parkinsons Disease by Using Machine Learning and Neural Network Methods, 2010 Second Int. Conf. Mach. Learn. Comput., pp. 218222, 2010.

[31] C.R.Pereiraetal.,Asteptowardstheautomateddiagnosisofparkinsons disease: Analyzing handwriting movements, Proc. - IEEE Symp. Comput. Med. Syst., vol. 2015July, pp. 171176, 2015.

[32] A. H. Al-fatlawi and M. H. Jabardi, Efficient Diagnosis System for Parkinson s Disease Using Deep Belief Network, pp. 13241330, 2016.

[33] Dahl, George E., Tara N. Sainath, and Geoffrey E. Hinton. ”Improving deep neural networks for LVCSR using rectified linear units and dropout.” Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on. IEEE, 2013.

[34] Srivastava, Nitish, et al. ”Dropout: a simple way to prevent neural networks from overfitting.” Journal of Machine Learning Research 15.1 (2014): 1929-1958.

[35] Bottou, Lon. ”Large-scale machine learning with stochastic gradient descent.” Proceedings of COMPSTAT’2010. Physica-Verlag HD, 2010. 177-186.

[36] Garraux, Gatan, et al. ”Multiclass classification of FDG PET scans for the distinction between Parkinson’s disease and a typical parkinsonian syndromes.” NeuroImage: Clinical 2 (2013): 883-893.

[37] Peng, Bo, et al. ”Computer aided analysis of cognitive disorder in patients with Parkinsonism using machine learning method with multilevel ROI-based features.” Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), International Congress on. IEEE, 2016.

[38] Rana, Bharti, et al. ”Graphtheorybased spectral feature selection for computer aided diagnosis of Parkinson’s disease using T1weighted MRI.” International Journal of Imaging Systems and Technology 25.3 (2015): 245-255.

[39] Billones, Ciprian D., et al. ”DemNet: A Convolutional Neural Network for the detection of Alzheimer’s Disease and Mild Cognitive Impairment.” Region 10 Conference (TENCON), 2016 IEEE. IEEE, 2016.

[40] Dou, Qi, et al. ”Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks.” IEEE transactions on medical imaging 35.5 (2016): 1182-1195.

[41] Abós, Alexandra, et al. "Discriminating cognitive status in Parkinson’s disease through functional connectomics and machine learning." Scientific reports 7 (2017): 45347.

[42] Kamagata, Koji, et al. "Connectome analysis with diffusion MRI in idiopathic Parkinson's disease: Evaluation using multi-shell, multi-tissue, constrained spherical deconvolution." *NeuroImage: Clinical* 17 (2018):518-529.