


Comparison Analysis of GLCM and PCA on Parkinson's Disease Using Structural MRI

Sanjana Tomer, J. C. Bose University of Science and Technology, YMCA, India

Ketna Khanna, J. C. Bose University of Science and Technology, YMCA, India

Sapna Gambhir, J. C. Bose University of Science and Technology, YMCA, India

 <https://orcid.org/0000-0001-5020-8000>

Mohit Gambhir, Innovation Cell, Ministry of Education, Faridabad, India

ABSTRACT

Parkinson's disease (PD) is a neurological disorder where the dopaminergic neurons experience deterioration. It is caused from the death of the dopamine neurons present in the substantia nigra (i.e., the mid part of the brain). The symptoms of this disease emerge slowly; the onset of the earlier stages shows some non-motor symptoms, and with time, motor symptoms can also be gauged. Parkinson's is incurable but can be treated to improve the condition of the sufferer. No definite method for diagnosing PD has been concluded yet. However, researchers have suggested their own framework out of which MRI gave better results and is also a non-invasive method. In this study, the MRI images are used for extracting the features. For performing the feature extraction techniques, gray level co-occurrence matrix and principal component analysis are performed and are analysed. Feature extraction reduces the dimensionality of data. It aims to reduce the feature of data by generating new features from the original one.

KEYWORDS

Feature Extraction, Gray Level Co-Occurrence Matrix, Magnetic Resonance Imaging, Parkinson's Disease, Parkinson's progression Markers Initiative, Principal Component Analysis, Statistical Parametric Mapping

INTRODUCTION

Disorders that mainly affect the central nervous system and the peripheral nervous system are called as neurological disorders. Parkinson disease (Anila M & Dr. G. Pradeepini, 2020) is one of the examples of this disorder and is second most common. PD begins in the middle or late in life. The exposure of pesticide (P. M. Shah et al., 2018) and Deoxyribonucleic acid (DNA) variation in genes like Leucine-rich repeat kinase 2 (LRRK2) can also cause the PD.

The substantia nigra (Aprajita Sharma & Mr. Ram Nivas Giri, 2013) helps to calibrate and fine tune the movement. It is a part of basal ganglia that controls the movement and connect to motor cortex. Substantia nigra is a latin term for black substance, it consists of two parts: pars reticula and pars compacta. Pars reticula receive signals from the striatum (which is also a part of basal ganglia)

DOI: 10.4018/IJIRR.289577

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

and relays messages to the thalamus via neurons rich in neurotransmitter Gamma aminobutyric acid (GABA). Pars compacta transmit messages to the striatum via neurons rich in neurotransmitter dopamine. This forms a nigrostriatal pathway, which help stimulate the cerebral cortex and initiate movement. Therefore, when substantia nigra pars compacta neurons, i.e., dopamine, die the hypokinetic state can be seen which is a common symptom (Tamanna Sood & Padmavati Khandnor, 2019) of PD. There are many motor and non-motor symptoms of PD that helps in detection of this disease. Some primary motor symptoms are tremors, rigidity, stooped posture, bradykinesia, hypokinesia, akinesia and postural instability. Whereas depression, dementia, sleep disturbances and anosmia are some of the non-motor symptoms which are caused by the dysfunction in dopaminergic signalling in other parts of brain. Various modalities have been proposed by the researchers to analyse these symptoms of PD such as Finger Tapping Test (FTT), handwritten test and voice samples. Some other neuroimaging technique such as: Magnetic Resonance Imaging (MRI), Computed Tomography (CT scan), Single Photon Emission Computed Tomography (SPECT), Positron Emission Tomography (PET), functional MRI (fMRI) are also suggested.

In FTT, the candidates are instructed to tap their index finger on the thumb fastly for a particular number of times. Initially the candidate is instructed to open their fingers widely and then their movements are analysed and the candidate is asked to stop. The score is generated on the 5-point scale.

Handwritten test as the name suggests, in this test the hand written datasets from PD. PD patients exhibit a reduction in amplitude of writing size. These patients also show slowness, irregular muscle contractions and tremors which leads to instability in patient's movements. The micrographia and impairments in writing helps in discrimination of a PD patient from a healthy individual. Therefore, when compared to healthy individuals, the PD patients apply less pressure, write small letters and spend more time acting. The patient's handwriting capability in language and their educational level will mostly differ from each other, for this reason the researchers focus on measuring the motor signs of hand writing while writing basic words, single letters or drawing Archimedean spirals.

Voice samples are used for detection of Parkinson's, though the voice modulation occurs gradually, the PD candidate can suffer from hypokinetic dysarthria. Collective or individual movements of jaws, lips and tongue are analysed (Pedro Gómez-Vilda et al., 2017). Various features like phonation, articulation, prosody and phonological are extracted for the classification of PD and HC.

PET mainly focusses on the neuroreceptors and pathways of neuronal transmission in various disease. It is an invasive method used for detecting the neurological disease in early stages. PET is done by injecting a tracer, a radioactive element, into the candidate's body. The tracer molecule emits the short-lived positron that emits the radiation and interact with the electron present in the candidate's body. The interaction of positron within the body is scanned by the PET scanners. These types of scanners examine the brain and note the time and place related to the radio-labelled molecules. 2-D and 3-D images of the brain can also be produced through PET scanner.

Computed Tomography scan uses the X rays for creating the 3-dimensional structure of the body. Blood vessels, bones and soft tissues can be shown through CT scan. It is invasive method, a special dye is either asked to consume orally or is injected into the candidate body for further analysis. After a span of time the candidate is gradually slide into the scanner. The x-ray machine start rotating and click loads of images for analysing the body part.

SPECT employs the usage of gamma rays. This technique provides true 3D information of the patient which can be formatted and controlled as required. This is similar to the traditional used planar imaging which used gamma camera.

In this technique, a gamma emitting radioisotope (tracer) is injected into patient's bloodstream and is attached to a particular ligand to form a radio gland. The tracer emits the gamma rays and the place of interest can be seen by the gamma camera.

MRI is a non-invasive method used to form the image of the anatomy of brain. The MRI help to detect the structural difference between the PD and HC brain. Through MRI, DIP can also be detected. In MRI, the terms T1 weighted, T2 weighted and FLAIR are the pulse sequences used to generate the

series of images. T1 and T2 weighted MRI are differentiated on the basis of two-time factors called time to echo and repetition time. In T1 weighted MRI there is short time to echo and short repetition time whereas both of these factors are long in T2 weighted MRI. The color of CSF in T1 weighting is dark and in T2 weighting it is bright. The gray matter appear bright in T1 weighting and is used for demonstrating the anatomy while in T2 weighted MRI the gray matter appear as light is helpful for the demonstration of pathology. In T1 and T2 scan, the brain appears lighter in the middle and darker in its surrounding. The opposite of this is visible in FLAIR the middle of the brain appears to be dark whereas its surroundings appear to be lighter. Fluid-attenuated inversion recovery (FLAIR) is a heavily T2-weighted technique that dampens ventricular CSF signals. In FLAIR, the gray matter appear to be as white, the white matter appears as gray and CSF look darker

The finger tapping test and handwritten test is conducted by the clinicians or researchers to analyse the motor skills of the candidate. However, various factors like age, hand dominance affect the results and can be misdiagnosed. The voice sample used for detection are also not so efficient. The neuroimaging techniques like CT, SPECT, PET and MRI are also used for the detection of PD. SPECT and PET are invasive methods i.e., radioactive tracers are injected into the body of the candidate opting for the test. Whereas in MRI no harmful substance is penetrated into the human body. CT scan take less time and is a lucid way to analyse the brain functioning. But it is less sensitive than MRI and can be performed if the candidate has any implanted device. MRI is a non-invasive method and require no pharmaceutical injections for conducting the test. It has high spatial resolution and good contrast. For this study, the MRI images are used as it give better results and is also a non-invasive method.

The main objective of this research is to focus on the feature extraction techniques proposed by the researchers to pull out interesting features from the data. It also aims to deeply study the functioning of brain by using neuroimaging techniques like MRI. The MRI dataset is collected from the PPMI database, i.e., Parkinson's Progression Markers Initiative organization. Through this various DICOM images are selected and analysed. After the rigorous understanding of the PD and its feature extraction techniques, the image processing has to be done.

Feature extraction (Martin D. Levine, 1969) is an imperative step of image pattern recognition (Hans J. Bremermann, 1968). The features extraction is a part of image processing. It plays an important role for the classification of data (Aqidatul Izza Poernama et al., 2019). The methods used for extracting the features may achieve better accuracy when the classification is performed. In general, the image processing can be achieved by initialising the data acquisition. The MRI dataset is selected from the Parkinson Progression Markers Initiative (PPMI) database (Ketna Khanna et al., 2020) and then the images are pre-processed. In the pre-processed stage various operations are performed on the images. The images are resized, reoriented and aligned in the initial phase of the preprocessing stage. The noise removal is done and then segmentation operations are performed.

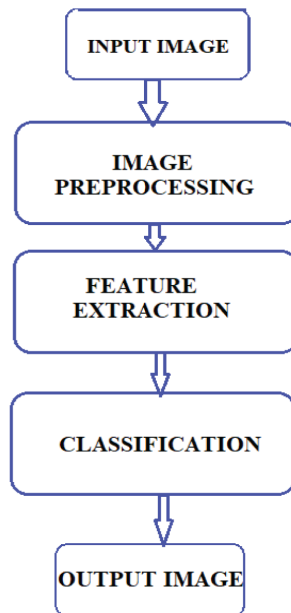
After image preprocessing, the original features are reduced to form a new set of features and there is no loss of crucial information. The feature extraction is achieved so that classification of the data can be done easily. Figure 1 shows the general flow of image processing.

LITERATURE REVIEW

Thus, various techniques have been proposed by the researchers for extracting the features of an image:

- Gabriel Solana-Lavalle & Roberto Rosas-Romero (2020) uses the voxel-based morphometry for the detection of the region of interest followed by the Haralick texture method and feature selection. This system is used for classifying the PD mri scans from the healthy individuals. The datasets used for the study was accessed from the PPMI database. Rather than focusing on a particular part of a brain, this work of author studies the whole brain. The research is conducted separately for both male and female. The detection performance for male group is 99.01% accuracy and for female is 97.97% accuracy.

Figure 1. Flow of image processing



- Ozkan Cigdem et al. (2019) used the Voxel-Based Morphometry (VBM) method for determining the morphological alterations between PD and Healthy Controls HC. Suggested two sample t test model i.e., 3D Gray Matter (GM) and White Matter (WM). By structural alteration, the 3D GM and WM tissue masks are obtained for t-contrast and f-contrast hypothesis. For selecting the volume of interest PCA and PDF (probability distribution function) are used. For comparing the performance of these methods SVM classifier is used. For GM the accuracy of PCA is 86.25% and for PDF is 75%.
- Gunjan Pahuja et al. (2016) introduces Extreme Learning Machine method with Genetics algorithm. The classifier uses the VBM features extracted from T1 weighted MRI. The MRI dataset is collected from the PPMI database, i.e., Parkinson's Progression Markers Initiative organization. The result of Genetic Algorithm- Extreme Learning Machine (GA-ELM) based classification approach is approximately 3% more accurate than Genetic Algorithm- Support Vector Machine (GA-SVM) classifier. The GA-ELM classifier has better performance. Ergo, it helps in better decision making and is helpful for medical practitioners in detecting PD. This approach can also be used in future with different modalities, like DT scan, for detecting different neuroimaging diseases.
- Li Zhang et al. (2016) used the machine learning framework based on PCA and SVM to achieve the automatic classification of PD and essential tremor (ET). The machine learning framework has two stages. At first, they used PCA to extract discriminative features from structural MRI data. Then SVM classifier is employed to classify PD and ET. And they employed statistical analysis method as an assisted tool to verify the differential brain regions.
- Shail Raval et al. (2020) have presented a comparative study for the early detection of Parkinson's. They provide the comparison analysis of symptoms like rigidity and bradykinesia. They have used SVM classifier, logistic regression, k-nearest neighbors (KNN) etc for the execution if the datasets. for detecting the tremor, the Random Forest Classifier (RF) on spiral test shows up to 99.79%, which is the highest of all the models and methods used.

- Priyanka Udayabhanu et al. (2016) proposes an automatic support system for the classification using Artificial neural networks (ANN). The decision making of the model is done in 2 stages, firstly the features are extracted using the GLCM method and at the second stage the classification is accomplished by using the Probabilistic Neural Network with Radial Basis Function (PNN-RBF) network. By using this 2 stages model better accuracy is obtained i.e., 90.9091%.
- Andrea Mechelli et al. (2004) uses the VBM and focus in the applications of healthy and diseased image. VBM helps in differentiating the structural brain differences and also helps to statistically analyse the volume tissues. It can also be used for analysing the gross structural abnormalities. This research presents the application of the disease subjects and healthy subject. Normalization and Segmentation Confounds, Pre-processing of Atypical Brains, Accuracy of Localization, Validity of Statistical Inferences, Interpretation, Sensitivity to Distributed and Non-linear Differences are some of the limitations discussed.
- Amirali Kazeminejad et al. (2017) investigated the appropriateness of graph theoretical approach with machine learning technique for diagnosing the PD. At first, rs-fMRI and graph theoretical approach is employed to study the brain of PD patients deeply. Then, both global and local graph measures are used. For this SVM classifiers, LOSO cross validation scheme and forward selection were used.
- Sara Soltaninejad et al. (2018) investigated the credibility of using MR images modality for detecting the PD. For obtaining the PD data set PPMI is accessed for processing the MR images. The FreeSurfer library is used for pre-processing and feature extraction which is the first and the second step of image processing. Classifiers like LR classifiers, RF classifiers and Support SVM are applied and their performance is compared. The system guarantees to detect the Parkinson sufferer.
- Beatrice Heim et al. (2017) summarizes the value of the different MRI techniques, such as Structural magnetic resonance imaging and Quantitative MRI, used for the detection of PD and review on the development of the parkinsonian syndrome. The defined automated methods use machine learning methods and support vector machine (SVM). The MRI analysis includes volumetric datasets, neuromelanin imaging, diffusion imaging and RS-functional MRI.
- Pir Masoom Shah et al. (2018) has suggested a Computer Aided Design based on Convolution Neural Network architecture which differentiate between PD and healthy control (HC). The three-layer convolutional layers are trained by the PPMI dataset which are used primarily. The T2-weighted MR images publicly available on PPMI are selected and aligned using image registration techniques. CNN as a classifier has an accuracy of 96% approx. It has better accuracy in comparison to SVM, RVM, decision tree, GA-ELM, ANN and other ML techniques.

However, the limitation of dataset was the major issue and it led to the overfitting of the CNN model but with the use of dropout layers in the network the issue of overfitting was avoided.

METHODOLOGY

Data Acquisition

The MRI dataset was obtained from the PPMI database i.e., Parkinson Progression Markers Initiative organization. The data of research groups 'PD' and 'NC' were used. The term PD here refers to Parkinson disease (Bharti Rana et al., 2015), this research group have MRI dataset of PD affected patients whereas NC stand for Normal Control, this group of individuals are healthy and have no neurological disease history. From both the research group 10 patients with T1 weighted MRI modality was selected. For performing the image pre-processing the original (or raw) image type were used in Digital Imaging and Communications in Medicine (DICOM) format. DICOM standard images helps in analysing the brain images. The DICOM images with .dcm extension stores the 2-dimensional

images of brain slices. It is a powerful and a flexible format used for storing the medical images. For a single subject hundreds of images are stored in 2D.

The data of research group PD and NC with age range of 55 to 65 was accessed from the PPMI database and both female and male participants were selected for doing this study. Table 1 shows the details of subjects used in this study.

MRI Data Pre-Processing

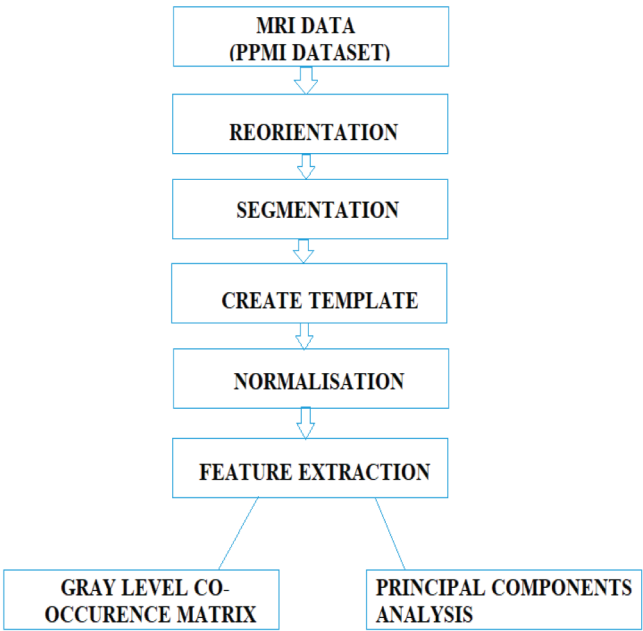
The pre-processing operations aims to improve the quality of the images by reducing or suppressing the unwilling image distortion and enhancing some image features which is suitable for segmentation. After that segmentation is perform to extract the real information data from input image in order to do further processing.

The MRI data (Zuliani Zulkoffli & Talha Afzal Shariff, 2019) is pre-processed by performing some operations like skull stripping, resizing, segmentation and normalization. Figure 2 depicts the sequence of work followed for performing the feature extraction. In this study, for performing the pre-processing of the data by Voxel Based Morphometry the package Statistical Parameter Mapping is used with the DARTEL toolbox. At the initial stage, the DICOM format images are converted into Neuroimaging Informatics Technology Initiative (NIFTI) by using the ‘DICOM import’ option available in Statistical Parametric Mapping SPM12.

Table 1. Details of subjects

Subjects	Age	Sex	Modality
PD	56-68	Both male and female	MRI
NC	55-60	Both male and female	MRI

Figure 2. Flow of work



These NIFTI files are saved in a different directory and then used for the reorientation and aligning the images, this stage helps in resizing and shaping the image. This process is followed by the segmentation, at this stage the brain tissues namely, white matter, gray matter and cerebrospinal fluid are separated. Although all the brain tissues are segmented, only Gray Matter segmentation is used while the WM and CSF are discarded. The Gray matter refers to the neurons and cells present in the central nervous system that are distributed in the region of cerebrum of the brain. The cerebrum consist of the thalamus, subthalamus and basal ganglia that are the most affected region in the PD.

After this normalization is done. For performing the spatial normalization, an original image is compared with a template image. For creating a template DARTEL toolbox is used which is available in SPM 12. DARTEL has precise intersubject alignment of MRI images and is also used for doing the normalization.

Then by performing the normalization the image is smoothed and the statistical analysis is done for calculating the volumes of the tissues.

Feature Extraction Techniques

Feature extraction is a method of reducing the size of data without losing the information from it. It is a method of processing the image by using the mathematical operations. The input can be image, series of image and sometimes it can be a video as well. There are many feature extraction techniques but here for the comparison analysis the GLCM and PCA (Bharti Rana et al., 2015) feature extraction techniques are used.

Before extracting the features, the pre-processed images are converted into jpg format using the MRIcro tool. It is an image viewing software used for analysing the medical images such as MRI, FMRI and PET. It helps in converting the .dcm and .nii format files to jpg, tiff and png so that they can be analysed easily. This software complements with the functioning of SPM software as well. Through this tool the neuroimagers can easily look up to all the slices and make it easy for extracting ROI and multiple ROI's as well. For this study, this tool is used for converting the pre-processed images obtained by SPM package. The nifti formats file were produced when the MRI data was pre-processed and for further procedure i.e., feature extraction, the medical images are converted into the analysed formats.

Gray Level Co-Occurrence Matrix (GLCM)

GLCM is a feature extraction technique. It is a statistical method used for reviewing the texture feature of an image (Happila.T & Kingston Stanley.P, 2015). The GLCM features are extracted from the gray image, and their intensities are then calculated for obtaining the spatial relationship to pixels by rows and columns. Various features like contrast, dissimilarity, asm etc can be obtained from GLCM.

GLCM features can be used for describing the pixel intensity of some part of the brain that are left at the pre-processed stage. The GLCM calculates the relation between the reference and the neighbour pixel i.e., it considers two pixels at an instance. Table 2 shows the feature of GLCM and its techniques (Aqidatul Izza Poernama et al., 2019).

Principal Component Analysis

Principal Component Analysis (Li Zhang et al., 2016) is a feature extraction technique used for reducing the dimensionality of the data. At this stage, the data is reduced in such a way that no crucial information from the data is lost. The data is reduced so that it become easy for classifying and visualising it. New features are derived from the existing features which have more information and have less redundant values.

PCA uses the orthogonal transformation to convert a set of correlated variables into the linearly uncorrelated variables called principal components (M. Shahbakhti et al., 2013).

Table 2. Feature of GLCM and its techniques

Features of GLCM	Equations
Contrast	$\sum_{ij}^{n-1} P(i-j)^2$
Correlation	$\sum_{ij}^{n-1} \frac{(i - \mu_x)(j - \mu_y)}{\sigma_x \sigma_y} P_{i,j}$
Entropy	$\sum_{ij}^{n-1} P_{ij} \log_2(P_{ij})$
Energy	$\sum_{ij}^{n-1} (P_{ij})^2$
Homogeneity	$\sum_{ij}^{n-1} \frac{P_{ij}}{1 + i - j }$

Thus, PCA elaborate the co-variance matrix i.e., the way a variable is associated with other variables. The PCA also explains the eigen values and eigen vectors. The eigen vectors represent the direction in which the data is spread whereas the eigen values represent the magnitude.

CLASSIFIERS

After extracting the features, the reduce image are then feed into the classifier so that the difference between PD and normal individuals can be done. For classifying the images, the random forest classifier (RF) is used for classifying between the two research groups. The Random Forest is a supervised learning algorithm that builds multiple decision trees and combine them to get stable and better prediction. The RF classifier produces more accurate and faster results than support vector machine. The RF classifier is performed by using the scikit-learn package which is tool used for predictive data analysis. The classification is implemented in jupyter notebook.

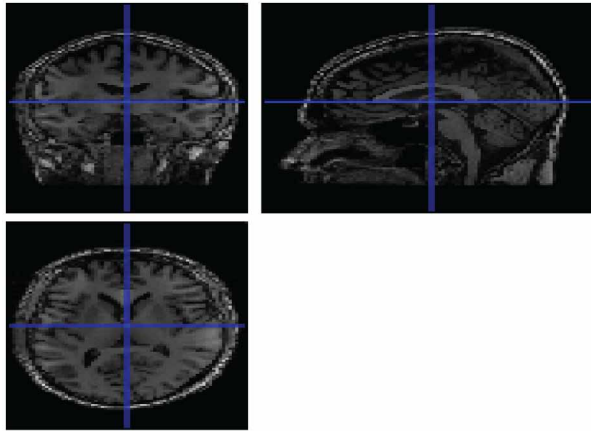
RESULTS AND DISCUSSIONS

Two different methods for extracting features were considered and implemented. The implementation was initialised by accessing the PPMI database and selecting the brain MRI images of specific research group. Then the images are pre-processed using the SPM toolbox.

SPM software is used for image pre-processing in MATLAB, it is a software designed for analysing the neuroimaging data. For this study SPM 12 version is used in MATLAB which was released on October 2014.

The final images obtained at this stage of image processing were used for performing the feature extraction techniques.

Figure 3. NIFTI images obtained after performing DICOM import



Pre-Processing

In SPM, the DICOM images selected from the PPMI database were converted into NIFTI format because SPM works with NIFTI standards. For converting the DICOM images into NIFTI, the 'DICOM import' function available in SPM package is used Figure 3 shows the output image obtained after performing the DICOM import which converts the DICOM images into NIFTI.

After the obtaining the .nii files, the images are aligned and reorientation are performed using SPM toolbox. The brain images are reoriented and the origin of the brain images is set to the anterior commissure. The brain images are resized and the motion correction is performed through SPM 12. The goal of motion correction is to adjust the series of images so that the brain is always in the same position. This minor step plays a crucial role for doing the segmentation and normalization without any error. Figure 4 shows the realigned and reoriented images.

In general, segmentation is used for removing background from the foreground. It helps in removing the noise from the data. Here the MRI images are then segmented in SPM 12 toolbox. This procedure is an extension of the algorithm of unified segmentation and is named as 'new segment' in SPM. By using this feature of SPM12 the brain images are skull stripped and separate brain tissues like GM, WM and CSF are obtained.

Figure 4. Images obtained after reorientation

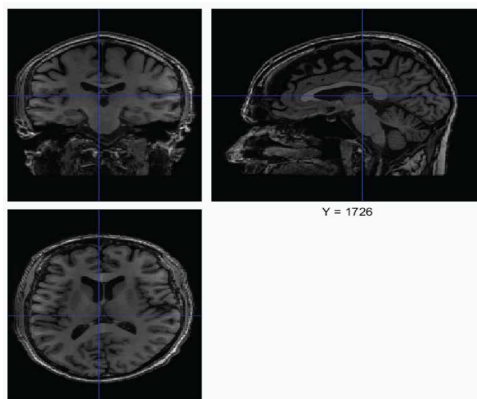
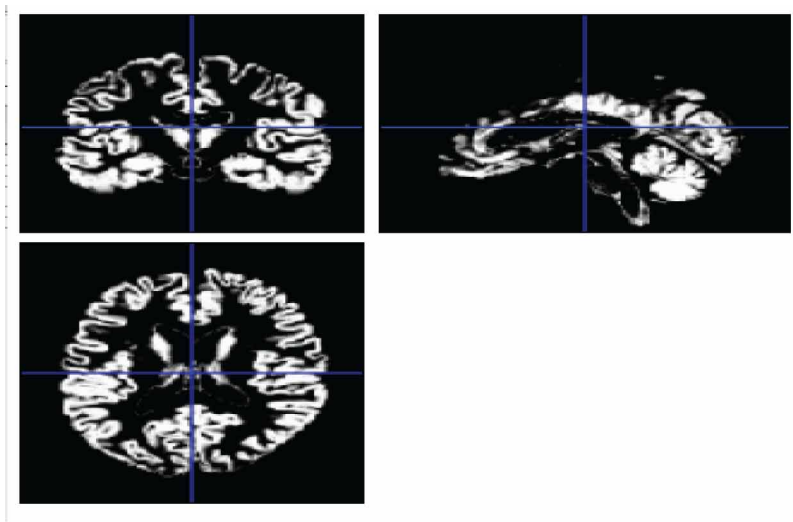


Figure 5. Segmented gray matter image



The aligned and reoriented images obtained previously were selected as an input for performing the segmentation and as an output segmented GM, WM and CSF images are obtained. From the segmented images, WM and CSF tissue images were discarded and only the GM segmented images were considered as the neuroimaging techniques have shown the severe loss of gray matter in Parkinson disease (Xia J, Miu J, Ding H, et al., 2013). Figure 5 shows the segmented image of gray matter.

For performing the normalization, a template is created. A template image is the brain with standard measurements and coordinates. In SPM, Dartel toolbox is used for creating the templates with .nii extension as shown in Figure 6.

Normalization is done by determining the similarities between the template image and subject image of an individual. The image normalization involves the continuous mapping i.e., done voxel by voxel between the template and the brain image of the participant.

Figure 6. Template image created by DARTEL toolbox

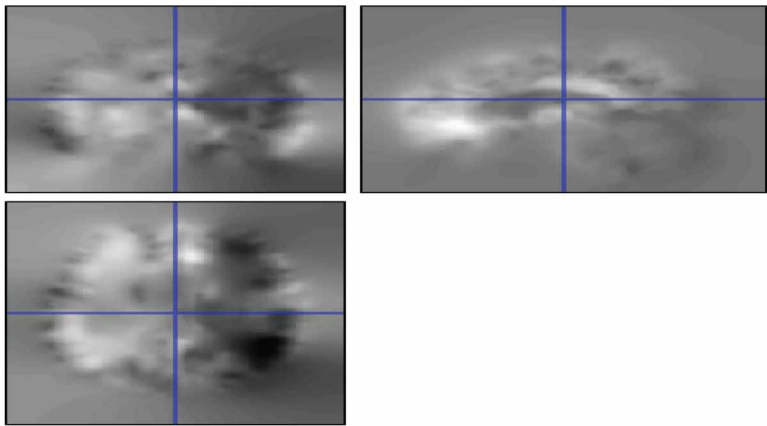
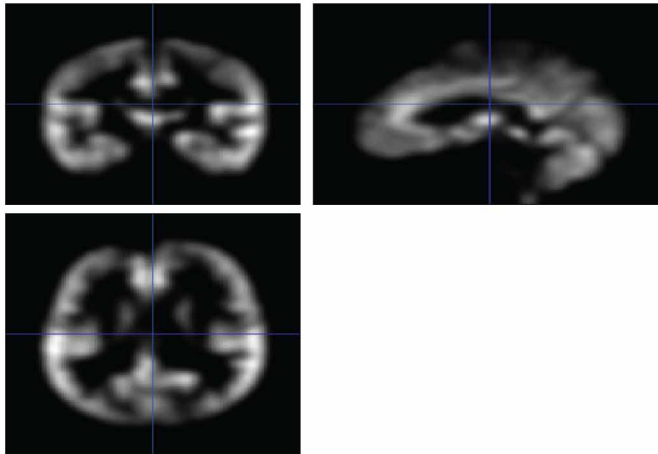


Figure 7. Normalised image



In SPM, the spatial normalization is performed by the DARTEL toolbox available in SPM. It determines the similarity between the original images and the templates.

Here, normalise to MNI space feature present in Dartel toolbox is used for performing the spatial normalization. By using this feature the smoothed images are generated in a way that the original signal is not lost. The parameter of full-width at half maximum (FWHM) of the Gaussian blurring is set to [8 8 8] which denotes the x, y and z direction. Figure 7 shows the normalised image.

After performing these steps, feature extraction has to be done for achieving the main aim of this study. The features are extracted using two methods i.e., GLCM and PCA.

Feature Extraction

Two methods were used for extracting the features from the pre-processed brain images. Gray Level Co-occurrence Matrix and Principal Components Analysis are two techniques used for reducing the dimensionality of the data.

In GLCM, the features were extracted from the gray level images. At an instance, two pixels were considered for calculating the relation between them. By using GLCM, 6 features were extracted based on different parameters. The GLCM method with distance 1 ($d=1$) and angle 0° , 45° , 60° and 90° were measured as shown in Figure 8.

Principal component analysis is used for reducing the dimensions of data. New features were created from the original features without any loss. The images were converted into the gray level for the feature extraction process. It is an unsupervised technique, which use mathematical principal to transfer a possibly correlated variables into number of variables called principal components.

In PCA, the images are transformed according to the components and the threshold value of cumulative variance is set accordingly. The measured cumulative explained variance is shown in Figure 9.

RESULTS

For this study, the performance of the model is calculated for both GLCM and PCA feature extraction techniques. The brain images of PD affected and healthy individuals were used from which various features were extracted for performing the GLCM feature extraction technique.

Figure 8. Features extracted using GLCM with Distance=1 (a) angle=0°, (b) angle=45°, (c) angle=60°, (d) angle=90°

	dissimilarity	correlation	homogeneity	contrast	ASM	energy		dissimilarity	correlation	homogeneity	contrast	ASM	energy
0	1.166667	-0.433861	0.450000	1.500000	0.277778	0.527046	0	1.25	-0.522233	0.475	2.25	0.250	0.500000
1	1.333333	-0.632456	0.433333	2.333333	0.277778	0.527046	1	0.50	0.870388	0.750	0.50	0.375	0.612372
2	1.333333	-0.447214	0.433333	2.333333	0.222222	0.471405	2	0.75	0.301511	0.625	0.75	0.250	0.500000
3	1.000000	-0.433013	0.566667	1.666667	0.222222	0.471405	3	1.00	-0.333333	0.600	2.00	0.375	0.612372
4	1.500000	-0.682288	0.350000	2.500000	0.222222	0.471405	4	0.75	1.000000	0.625	0.75	0.375	0.612372
...
843	0.833333	0.000000	0.583333	0.833333	0.333333	0.577350	843	0.50	0.333333	0.750	0.50	0.375	0.612372
844	0.833333	-0.420084	0.616667	1.166667	0.277778	0.527046	844	1.50	-0.577350	0.400	3.00	0.375	0.612372
845	1.166667	-0.585540	0.483333	1.833333	0.222222	0.471405	845	1.50	-0.707107	0.350	2.50	0.250	0.500000
846	0.500000	0.581318	0.783333	0.833333	0.222222	0.471405	846	1.00	-0.454545	0.600	2.00	0.250	0.500000
847	1.000000	-0.415227	0.566667	1.666667	0.277778	0.527046	847	0.75	-0.577350	0.625	0.75	0.375	0.612372

(a)

	dissimilarity	correlation	homogeneity	contrast	ASM	energy
0	1.666667	-0.862069	0.300000	3.000000	0.277778	0.527046
1	0.833333	0.000000	0.616667	1.166667	0.277778	0.527046
2	1.500000	-0.664364	0.350000	2.500000	0.222222	0.471405
3	1.166667	-0.297044	0.450000	1.500000	0.277778	0.527046
4	0.500000	0.447214	0.750000	0.500000	0.222222	0.471405
...
843	1.000000	0.000000	0.600000	2.000000	0.277778	0.527046
844	1.333333	-0.585491	0.400000	2.000000	0.222222	0.471405
845	0.833333	-0.316228	0.650000	1.500000	0.388889	0.623610
846	1.000000	-0.034483	0.566667	1.666667	0.222222	0.471405
847	0.333333	0.333333	0.833333	0.333333	0.277778	0.527046

(b)

	dissimilarity	correlation	homogeneity	contrast	ASM	energy
0	1.500000	-0.656532	0.383333	2.833333	0.277778	0.527046
1	0.666667	-0.759257	0.700000	1.000000	0.388889	0.623610
2	1.000000	-0.273861	0.566667	1.666667	0.222222	0.471405
3	1.500000	-0.682288	0.350000	2.500000	0.277778	0.527046
4	0.833333	-0.216930	0.650000	1.500000	0.277778	0.527046
...
843	0.666667	-0.158114	0.700000	1.000000	0.222222	0.471405
844	1.000000	-0.200000	0.533333	1.333333	0.222222	0.471405
845	0.666667	0.000000	0.700000	1.000000	0.222222	0.471405
846	1.166667	-0.415227	0.516667	2.166667	0.277778	0.527046
847	0.666667	0.100000	0.700000	1.000000	0.222222	0.471405

(c)

	dissimilarity	correlation	homogeneity	contrast	ASM	energy
0	1.666667	-0.862069	0.300000	3.000000	0.277778	0.527046
1	0.833333	0.000000	0.616667	1.166667	0.277778	0.527046
2	1.500000	-0.664364	0.350000	2.500000	0.222222	0.471405
3	1.166667	-0.297044	0.450000	1.500000	0.277778	0.527046
4	0.500000	0.447214	0.750000	0.500000	0.222222	0.471405
...
843	1.000000	0.000000	0.600000	2.000000	0.277778	0.527046
844	1.333333	-0.585491	0.400000	2.000000	0.222222	0.471405
845	0.833333	-0.316228	0.650000	1.500000	0.388889	0.623610
846	1.000000	-0.034483	0.566667	1.666667	0.222222	0.471405
847	0.333333	0.333333	0.833333	0.333333	0.277778	0.527046

(d)

	dissimilarity	correlation	homogeneity	contrast	ASM	energy
0	1.500000	-0.656532	0.383333	2.833333	0.277778	0.527046
1	0.666667	-0.759257	0.700000	1.000000	0.388889	0.623610
2	1.000000	-0.273861	0.566667	1.666667	0.222222	0.471405
3	1.500000	-0.682288	0.350000	2.500000	0.277778	0.527046
4	0.833333	-0.216930	0.650000	1.500000	0.277778	0.527046
...
843	0.666667	-0.158114	0.700000	1.000000	0.222222	0.471405
844	1.000000	-0.200000	0.533333	1.333333	0.222222	0.471405
845	0.666667	0.000000	0.700000	1.000000	0.222222	0.471405
846	1.166667	-0.415227	0.516667	2.166667	0.277778	0.527046
847	0.666667	0.100000	0.700000	1.000000	0.222222	0.471405

Figure 9. Cumulative explained variance obtained using PCA

	variance	cumulative variance
0	2	[81.32724949671868, 91.4515881967242, 96.54871...
1	2	[84.03136563426771, 94.33179433725685, 98.1319...
2	1	[86.23244996765132, 97.05613006751854, 99.2226...
3	1	[88.37910332526474, 98.41815933811742, 99.6052...
4	1	[91.65622744686989, 99.09668238536993, 99.7188...
...
843	0	[96.16147072398971, 98.93567308733475, 99.6455...
844	0	[96.4846173245298, 99.09585223167863, 99.71635...
845	0	[96.75852527199463, 99.21118327850398, 99.7280...
846	0	[96.94254368959957, 99.16950076398874, 99.6136...
847	0	[97.01750499353822, 99.20527782296813, 99.4256...

GLCM techniques is implemented for the distance value 1 and four different angle value. The correctness of the model for distance (d=1) angle (0°), distance (d=1) angle (45°), distance (d=1) angle (60°) and distance (d=1) angle (90°) are measured as shown in table 3.

GLCM features accuracy for angle 0° and 90° measure the same whereas its scored the lowest with 45° angle measurement.

The same dataset was used for extracting the features from the brain images by using the PCA and the accuracy of this model was calculated by using the RF classifier, same as performed with the former feature extraction technique, the model scored only 87.5%.

Table 3. The performance of GLCM method in different degrees with neighbourhood distance 1 and performance of PCA method

Technique	Accuracy
GLCM [0]	90.5
GLCM [45]	87
GLCM [60]	90.5
GLCM [90]	89.5
PCA	87.5

CONCLUSION

For the comparison analysis of two different feature extraction techniques, MRI brain images were selected. The GLCM and PCA were chosen for extracting the features from the brain images accessed by the PPMI database. In total 20 individual's data of PD and NC group were selected. Viz, 968 pre-processed images were accessed from which only 848 were used and the left ones were discarded.

The GLCM method is used for extracting the features. The texture features are extracted from the gray image by using the gray level co-occurrence matrix. Various features like contrast, dissimilarity and asm were be extracted. The matrices measure the relationship between the pixels. The GLCM method parameters like distance, levels and degree of angles were changed for analysing the performance of the method. For every distance value $d=1$, angle (0° , 45° , 60° and 90°) values were measured and it can be analysed that this method is sensitive to any change in occurring image data. After performing GLCM method, it can be concluded that it consumes less computational time and the output is obtained in short time only.

The Principal component analysis is used for extracting the feature for analysing the PD affected brain images. For performing the PCA, the images are converted into Gray level for reducing the dimensions of the data. PCA finds the eigenvector from the co variance matrix from which is then used for reducing the feature data.

After performing the PCA method practically, it can be concluded that the PCA, in comparison of GLCM, consumes high computational time and requires large memory size for performing the feature extraction method.

The GLCM method gave more accurate results in comparison to the PCA method. The GLCM method with distance ($d=1$) and angle (0° and 90°) scored 90.5% accuracy whereas PCA method scored only 87.5%.

REFERENCES

- Anila, M. (2020). A Review on Parkinson's Disease Diagnosis using Machine Learning Techniques. *International Journal of Engineering Research & Technology (Ahmedabad)*, 9(6), 2272–2278. doi:10.17577/IJERTV9IS060241
- Bremermann, H. J. (1968). Pattern Recognition, Functionals, and Entropy. *IEEE Transactions on Biomedical Engineering, BME-15*(3), 201–207. doi:10.1109/TBME.1968.4502565 PMID:5667807
- Cigdem, O., Demirel, H., & Unay, D. (2019). The Performance of Local-Learning Based Clustering Feature Selection Method on the Diagnosis of Parkinson's Disease Using Structural MRI. *IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 1286-1291. doi:10.1109/SMC.2019.8914611
- Gómez-Vilda, P., Mekyska, J., & Ferrández, J. M. (2017). Parkinson Disease Detection from Speech Articulation. *Neuromechanics*.
- Happila, T., & Kingston Stanley, P. (2015). Brain Pattern Recognition Based Classification of Neurodegenerative Diseases. *2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIECS)*, 1-5.
- Heim, B., Krismer, F., De Marzi, R., & Seppi, K. (2017). Magnetic resonance imaging for the diagnosis of Parkinson's disease. *Journal of Neural Transmission (Vienna, Austria)*, 124(8), 915–964. doi:10.1007/s00702-017-1717-8 PMID:28378231
- Kazeminejad, A., Golbabaei, S., & Soltanian-Zadeh, H. (2017). Graph Theoretical Metrics and Machine Learning for Diagnosis of Parkinson's Disease Using rs-fMRI. *2017 Artificial Intelligence and Signal Processing Conference (AISP)*, 134-139. doi:10.1109/AISP.2017.8324124
- Khanna, K., Gambhir, S., & Gambhir, M. (2020). Article. *Journal of Critical Review*, 7(18), 1461–1467.
- Levine, M. D. (1969). Feature Extraction: A Survey. *Proceedings of the IEEE*, 57(8), 1391–1407. doi:10.1109/PROC.1969.7277
- Mechelli, A., Price, C. J., Friston, K. J., & Ashburner, J. (2004). Voxel-Based Morphometry of the Human Brain: Methods and Applications. *Current Medical Imaging*, 1(1), 1–9.
- Pahuja, G., & Nagabhushan, T. N. (2016). A Novel GA-ELM approach for Parkinson's disease detection using Brain Structural T1-weighted MRI Data. *Second International Conference on Cognitive Computing and Information Processing (CCIP)*, 1-6. doi:10.1109/CCIP.2016.7802848
- Poernama, A. I., Soesanti, I., & Wahyunggoro, O. (2019). Feature Extraction and Feature Selection Methods in Classification of Brain MRI Images: A Review. *International Biomedical Instrumentation and Technology Conference (IBITeC)*, 58-63. doi:10.1109/IBITeC46597.2019.9091724
- Poernama, A. I., Soesanti, I., & Wahyunggoro, O. (2019). Feature Extraction and Feature Selection Methods in Classification of Brain MRI Images: A Review. *2019 International Biomedical Instrumentation and Technology Conference (IBITeC)*, 58-63. doi:10.1109/IBITeC46597.2019.9091724
- Rana, B., Juneja, A., Saxena, M., Gudwani, S., Kumaran, S. S., Behari, M., & Agrawal, R. K. (2015). Graph-Theory-Based Spectral Feature Selection for Computer Aided Diagnosis of Parkinson's Disease Using T1-Weighted MRI. *International Journal of Imaging Systems and Technology*, 25(3), 245–255. doi:10.1002/ima.22141
- Rana, B., Juneja, A., Saxena, M., Gudwani, S., Senthil Kumaran, S., Agrawal, R. K., & Behari, M. (2015). Regions-of-interest based automated diagnosis of Parkinson's disease using T1-weighted MRI. *Expert Systems with Applications*, 42(9), 4506–4516. doi:10.1016/j.eswa.2015.01.062
- Raval, S., Balar, R., & Patel, V. (2020). A Comparative Study of Early Detection of Parkinson's Disease using Machine Learning Techniques. *2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)*(48184), 509-516.
- Shah, Zeb, Shafi, Zaidi, & Shah. (2018). Detection of Parkinson's Disease in Brain MRI using Convolutional Neural Network. In *Proceedings of the 24th International Conference on Automation & Computing* (pp. 1-6) Newcastle University.

- Shah, Zeb, Shafi, Zaidi, & Shah. (2018). Detection of Parkinson Disease in Brain MRI using Convolutional Neural Network. In *Proceedings of the 24th International Conference on Automation & Computing*. Newcastle University. doi:10.23919/IConAC.2018.8749023
- Shahbakhti, M., Taherifar, D., & Zareei, Z. (2013). Combination of PCA and SVM for diagnosis of Parkinson's disease. *2013 2nd International Conference on Advances in Biomedical Engineering*, 137-140.
- Sharma, & Giri. (2013). An Elegant Approach for Diagnosis of Parkinson's disease on MRI Brain Images by Means of a Neural Network, *Ijesrt. International Journal of Engineering Sciences & Research Technology*, 2553–2557.
- Solana-Lavalle, G., & Rosas-Romero, R. (2020). Classification of PPMI MRI scans with voxel-based morphometry and machine learning to assist in the diagnosis of Parkinson's disease. *Computer Methods and Programs in Biomedicine*, 198, 1–15. PMID:33099263
- Soltaninejad, S., Cheng, I., & Basu, A. (2018). Towards the Identification of Parkinson's Disease Using only T1 MR Images. *International Conference on Smart Multimedia*, 145–156. doi:10.1007/978-3-030-04375-9_13
- Sood, T., & Khandnor, P. (2019). Classification of Parkinson's Disease Using Various Machine Learning Techniques. *International Conference on Advances in Computing and Data Sciences*, 1045, 296–311. doi:10.1007/978-981-13-9939-8_27
- Udayabhanu, Viswan, & Padmarajan. (2016). MRI Brain Image Classification using GLCM Feature Extraction and Probabilistic Neural Networks. *IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)*, 4-15.
- Xia, J., Miu, J., Ding, H., Wang, X., Chen, H., Wang, J., Wu, J., Zhao, J., Huang, H., & Tian, W. (2013). Changes of brain gray matter structure in Parkinson's disease patients with dementia. *Neural Regeneration Research*, 8(14), 1276–1285. PMID:25206422
- Zhang, L., Liu, C., & Zhang, X. (2016) Classification of Parkinson's disease and Essential Tremor Based on Structural MRI. *10th International Conference on Software, Knowledge, Information Management & Applications (SKIMA)*, 410-412.
- Zhang, L., Liu, C., Zhang, X., & Tang, Y. (2016). Classification of Parkinson's disease and Essential Tremor Based on Structural MRI. *7th International Conference on Cloud Computing and Big Data*, 353-356.
- Zulkoffli, Z., & Shariff, T. A. (2019). Detection of Brain Tumor and Extraction of Features in MRI Images Using K-means Clustering and Morphological Operations. *IEEE International Conference on Automatic Control and Intelligent Systems*, 1-5. doi:10.1109/I2CACIS.2019.8825094

Ketna Khanna is currently pursuing Ph. D. in Computer Engineering from J.C. Bose University of Science and Technology, YMCA, Faridabad. She is Master of Technology in Computer Engineering and Bachelor of Technology in Computer Science & Engineering. Her main research area focuses on Machine Learning, Data Mining and Software Testing.

Sapna Gambhir, IEEE senior member and treasurer WIE AG IEEE Delhi section, is an associate Professor of Computer Engineering Department in J C Bose University of Science & Technology, YMCA, Faridabad (India). She has completed her Doctorate in Computer Engineering from Jamia Millia Islamia, Delhi in 2010. She has approx. 20 years of experience in academics and research & development. She has been awarded with "Best Woman Scientist Award" under Research Excellence and Academic Awards (REAA-2018) and "Best Teacher Award" from District Commissioner of Faridabad. She has guided many Ph.D students in the area of wireless networks, Internet of Things, Mobile Adhoc Networks and medical imaging. She has published more than 80 research articles in peer-reviewed journal and conference. Two patents, out of which one is with USPTO and One is in Indian Patent Office (IPO) has been granted. She has already guided three Ph.Ds successfully as sole supervisor. Currently, three Ph.D scholars are doing their research under her guidance. She has also participated in UKIERI 2019 leadership program which is a collaborative initiative by Govt. of UK and India. She is a principal investigator of research projects of many government organizations.