Identifying shape-based biomarkers for diagnosis of Parkinson's Disease from Ioflupane (^{123}I) SPECT data

Vanessa Azzopardi[1], Dr. M. Guy[2] and Dr. E. Lewis[12]

[1] University of Surrey, 388 Stag Hill, Guildford GU2 7XH, UK
[2] University Hospital Southampton, Tremona Road, Southampton SO16 6YD, UK
vanessa.azzopardi2@gmail.com
emma.lewis2@uhs.nhs.uk
matthew.guy@uhs.nhs.uk

Abstract. Parkinson's disease is a progressive, incurable neurodegenerative condition affecting movement which has been linked to poor quality of living and considerable socio-economic burdens. To date, treatment can at best slow down the degradation process. However, successful disease management is subject to early detection of the disease which, in turn, depends on the diagnostic process.

Clinical investigation alone has proven insufficient in discriminating between early Parkinson's disease and essential tremor. Functional neuroimaging circumvents this problem by visualising dopamine transporter concentrations in the brain, providing a differential even during the early stages of the disease. Yet the traditional visual assessment of SPECT data introduces subjectivity and susceptibility to variation whilst being impractical for monitoring and assessing disease progression.

This work, presents a machine-learning approach to the assessment of three-dimensional SPECT data. The system extracts intensity and shape information from the data following binarisation which utilises an experimental approach towards the identification of an optimal threshold. The striatal binding ratio is calculated based on the three-dimensional data rather than two-dimensional clinical standard. The resulting semi-quantitative measure and the extracted intensity and shape information are collectively used as data features and are subjected to a support vector machine to classify between positive and negative cases of Parkinson's disease. The classification system is reported to attain an average accuracy of 97%; with 96.6% sensitivity and 97.8% specificity. This shows an improvement over the clinical standard visual assessment which reportedly attained 94% sensitivity and 92% specificity.

Keywords: Classification, Machine Learning, Computer-Aided Diagnosis, Parkinson's Disease, Single-Photon Emission Computed Tomography.

1 Introduction

1.1 Parkinson's Disease

Parkinson's disease (PD) is characterized by loss of dopamine transporters (DaT) particularly in the striatum, a key region in motor control [1]. Typical symptoms of PD include: resting tremor, bradykinesia (slow movement), rigidity and postural instability, all of which are associated with considerable morbidity and socio-economic burdens [2]. To date, PD is incurable and the degenerative nature of the disorder implies that the damage inflicted by the disease increases throughout life. Thus, the earlier the disease is detected, the higher the possibility of preserving life [3]. The diagnosis route is lengthy, often involving a clinical assessment followed by the analysis of the response to medication, such as levodopa, as well as nuclear imaging scans.

1.2 Clinical Background

Research has established that clinical examinations often prove to be incorrect after subsequent follow-ups [4]. In many cases PD is mistaken for a distinct condition with similar symptoms, known as Essential Tremor [5]. Disease management differs greatly between the two as does prognosis and hence inaccurate diagnosis leads to under- or over-treatment with medication being prescribed unnecessarily [6, 7].

Typically, PD is confirmed upon reduction in the severity of PD symptoms following periodic administration of levodopa [8]. However, in the early stages of the disease, symptoms tend to be mild and atypical with a poor response to levodopa, making for complicated diagnosis, often with inconclusive results [9-11].

Despite this, research suggests that the neurological signs of PD are present long before the manifestation of clinical symptoms with at least 50% striatal dopamine loss taking place prior to the emergence of clear clinical symptoms [2]. This motivates the application of single-photon emission computed tomography (SPECT) in order to visualize DaT concentrations in the brain [12]. DaT are known to be partially lacking in early PD cases but are often found in healthy concentrations in the presence of essential tremor. As a result, it has been revealed that the use of SPECT imaging has enabled the differentiation of PD patients from healthy subjects and patients with essential tremor with approximately 90% sensitivity and specificity [13] [14] [15].

1.3 DaT SPECT Imaging for Parkinson's Disease Diagnosis

DaT SPECT imaging results in greyscale images where the intensity of each pixel is directly correlated with the counts registered by the gamma camera. Areas of high intensity indicate a high DaT concentration. Data is typically acquired in three-dimensions comprising a volume which is visually inspected by a clinician. The observer seeks a homogeneous, symmetrical 'comma-shaped' pattern which depicts healthy striatal uptake [16]. Any other pattern is classified as abnormal with early PD cases typi-

cally demonstrating reduced uptake in the putamen and hence a 'full stop-shaped' pattern [2]. Being a visual assessment, this approach is inherently subject to inter-observer variability as well as human error.

1.4 Project Outline

This work presents an automatic classification system which differentiates between positive and negative cases of PD. DaT SPECT volumes are analysed in terms of shape, intensity and semi-quantitative measures. These are used to form features representing the said volumes during classification which is carried out using a support vector machine (SVM).

The images resulting from SPECT provide a qualitative measure of the dopaminergic system integrity from a direct correlation of the image context with neurological function. However, this falls short of the requirements for long-term assessment of disease progression due to the lack of quantitative information. Presently, the striatal binding ratio (SBR) is occasionally used at some imaging centres in order to provide quantitative information, however, this is generally calculated from a two-dimensional image, neglecting considerable information [17]. In this work, efforts were directed towards making the most out of the 3D data available; including, among other things, the extraction of a three-dimensional version of the SBR. All features are extracted following consideration of the entire volume and this also applies to the SBR measure used in this work.

2 Study Dataset

The data used in this study was provided by the Parkinson's Progression Markers Initiative (PPMI) under which a large-scale study was conducted to acquire clinical data from individuals of significant interest with respect to PD. The collected data was reviewed by experts and compiled into a publicly-available database and biorepository for research purposes [18]. The work presented in this document is based on a portion of this data. It comprises DaTScan SPECT volumes acquired longitudinally and uniformly from 344 early Parkinson's disease patients and 193 healthy control subjects [19].

SPECT imaging scans were carried out using 123I-FP-CIT (also known as DaTScan) administered in doses of 185MBq and imaged 4 ± 0.5 hours later [20]. Raw projection data was acquired into a 128 x 128 matrix with a stepping angle of 3° and with the acquisition window centred on $159 \pm 10\%$ keV. A parallel hole collimator was used in step and shoot acquisition mode with each head rotating 360°.

Pre-processing involved iterative reconstruction using ordered subset expectation maximization (OSEM) through the HOSEM reconstruction program within the HERMES package. Attenuation correction and filtering were carried out using the Chang 0 method and a standard Gaussian 3D 6mm filter respectively. Images were anatomically aligned (registered) and normalized to the standard Montreal Neurologic Institute (MNI) space [9]. The resulting scan volumes were stored in DICOM format.

3 Methodology

3.1 Volume Visualisation

The 3D DaT SPECT volumes acquired from the PPMI repository were first visualised with the help of a tailor-made interactive console. Using this console, the user is allowed to navigate through the 3D volume by manually placing a pair of crosshairs at any location within the volume and view the three cardinal image slices (transverse, sagittal and coronal views) along with the corresponding number of registered counts at that point. This approach enabled an understanding of the dataset and the existing visual differences between the healthy and pathological subsets. Fig. 1 illustrates the mentioned console being used to view the control data.

3.2 Data Cropping

The initial visualisation step led to the observation of radiopharmaceutical uptake towards the jaw region of the head.

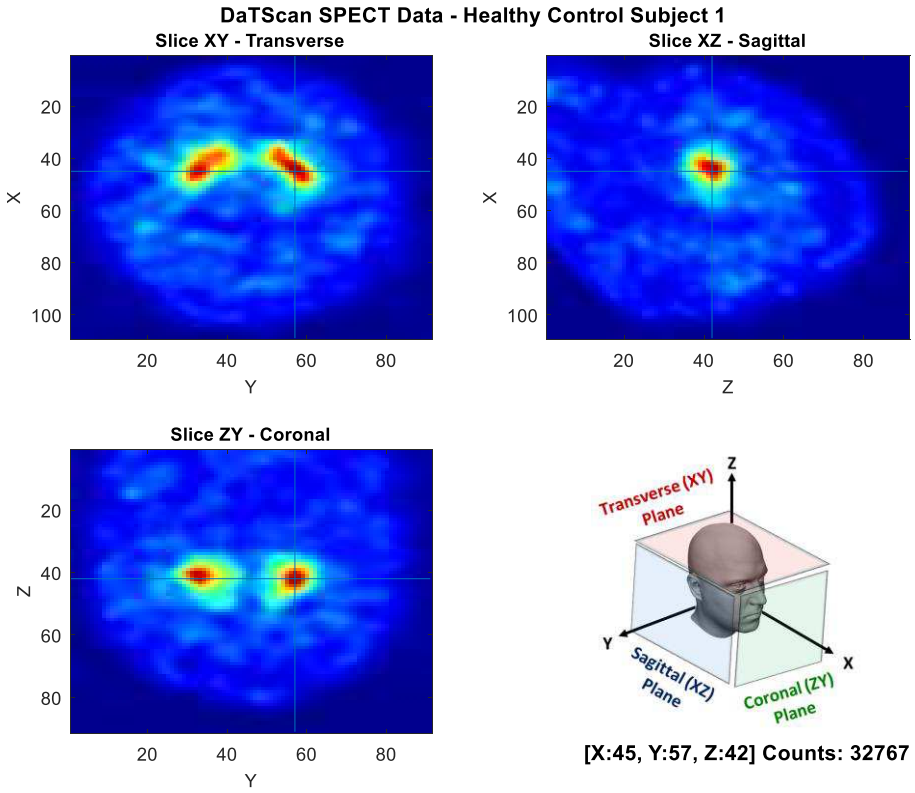

Fig. 1. The slice viewer console displaying transverse, sagittal and coronal image slices from the DaT SPECT data of a healthy control subject.

This is a known phenomenon whereby the parotid glands take up the radiopharmaceutical and therefore show up as regions of high activity in the DaT SPECT volume. As a result, locating the striatal uptake region by thresholding the voxel count risks confusion with the striatal regions which are the true markers of DaT presence in the brain. To prevent this, the 3D data was restricted to a smaller 60x40x25 pixel volume which captured the entire striatal region without incorporating uptake corresponding to regions outside the anatomical region of interest.

The dimensions of the mentioned volume were determined following manual testing. This involved plotting an arbitrary cuboidal boundary onto the 3D visualisations. The initial dimensions of the cuboid were selected to incorporate the striatal region. The 3D DaT SPECT visualisations were programmed to display sequentially for the entire dataset. Regions of striatal uptake lying outside of the plotted boundary were identified and triggered a correction in the dimensions of the cuboidal boundary. The procedure was repeated following each alteration in dimensions. This manual testing process converged, resulting in the optimal volume having the dimensions 60x40x25 pixels. This testing was based on a dataset of 400 DaT SPECT 3D volumes.

The defined 60x40x25 pixel volume was centred on the centroid of striatal edge pixels. Edge detection was carried out on the 2D transverse slice which demonstrated the striatal uptake region most clearly. By default, this corresponded to the 2D transverse slice which housed the voxel registering the highest count, that is, the best slice.

Each best slice was sharpened using unsharp masking and subsequently low-pass filtered [21]. This process resulted in the suppression of background artefacts and the accentuation of edges within the best slice, facilitating edge detection. Fig. 2 illustrates samples of the best slice at the different stages of this process.

Edge detection was carried out through convolution by Sobel operators. The resulting binary image of the best slice consisted of edge pixels from both the striatal region and the boundary of the patient's head, as shown in Fig. 3A.

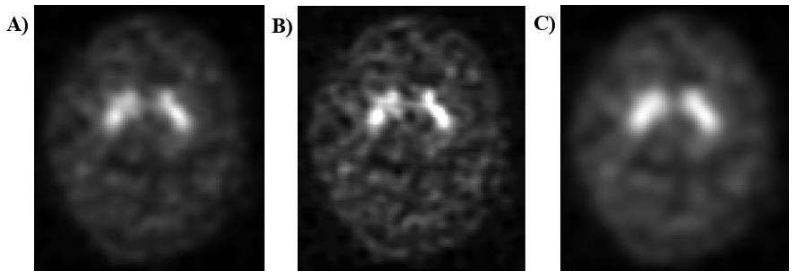

Fig. 2. The original best slice (A), the sharpened best slice (B) and the sharpened and blurred best slice (C) for a healthy control subject.

In order to remove unwanted artefacts, the best slice was multiplied by the circular binary mask shown in Fig. 3C. The circular binary mask was centered at the centroid of all edge pixels in the image [21]. The radius of the mask was set to 30 pixels, following iterative testing. Fig. 3B depicts the circumference of the circular binary mask superimposed on the best slice for illustration purposes. Fig. 3D shows the resulting

clean representation of striatal edge pixels in the best slice following the multiplication described. This procedure proved successful across the entire dataset due to the spatial alignment and co-registration of the images [22] [18] [19] [23].

At this point, the best slice was a binary image consisting of edge pixels which pertain to the striatal region. The centroid of these pixels was computed and used as a 3D reference point. A 60x40x25 pixel volume centered upon this reference point was cropped out of 3D DaT SEPCT volume for each patient. Fig. 3E provides a 2D illustration of this volume centering procedure. The area within the red box represents the new cropped 3D volume. This procedure led to the reduction of the DaT SPECT volume of each patient to a smaller region which excludes any high-activity regions which are unrelated to the striatum. The dimensions of this volume of interest were determined empirically, following iterative visual testing. This was carried out on roughly 75% of the data (400 subjects) selected at random. These included both Parkinson's disease patients and healthy controls. Having a generous dataset enabled the selection of parameters which suit a wider variety of cases. This procedure however, preserved a quarter of the dataset to avoid overfitting and allow for fair testing and classification at later stages.

3.3 Segmentation

Each 3D volume was segmented in order to separate the volume of interest (the striatal region) from the surrounding volume. Segmentation was achieved by thresholding the

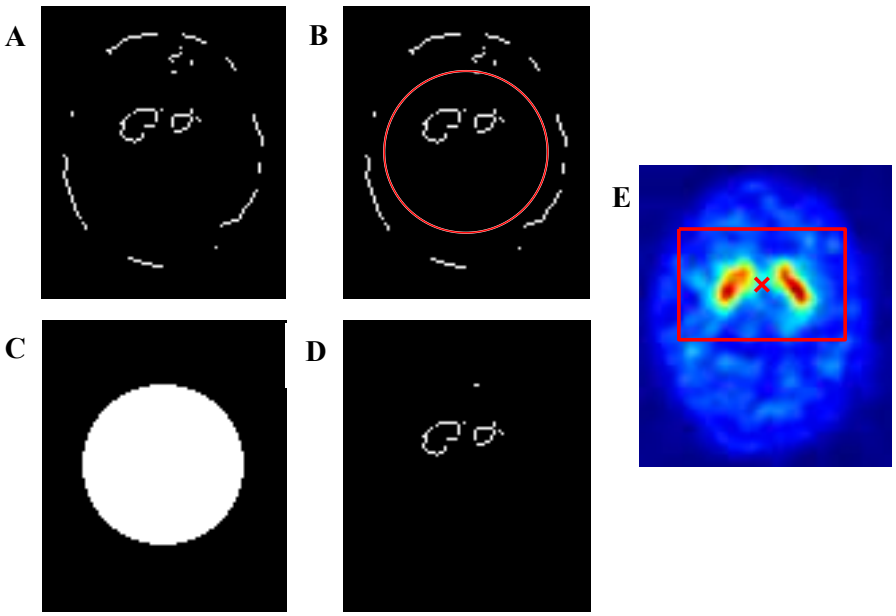

Fig. 3. The original edge pixels of the best slice (A), the circle separating striatal edge pixels from the head boundary edge pixels (B), the circular binary mask (C) and the striatal edge pixels (D). A 2D illustration of the volume centering procedure (E).

registered counts within the 3D volume according to Eqn. 1. This computation is based on the mean intensity or voxel value of the 3D volume (μ) and standard deviation of the intensity or voxel values within the 3D volume (σ).

$$threshold = 0.6 - (\mu(0.3 - \mu)) + \frac{\sigma}{2} \quad (1)$$

The 3D data of each subject was initially normalised to the maximum intensity (voxel value) of each 3D volume. This resulted in intensity values between 0 and 1, creating a common standard across the dataset. A comparison of values across the subject groups revealed that control data resulted in a lower mean intensity when compared to pathological data. As a result, the simple numerical threshold of 0.6 which is suggested by Lyra et al. and Prashant et al., proved insufficient [24] [25]. For this reason, the mean intensity (μ) was factored into the equation.

Thresholding was applied using a threshold equation which simply subtracted the normalised mean intensity from 0.6. Visual validation of the data following thresholding indicated that the larger the difference between the mean and 0.6, the larger the required increase in the threshold value, hence the need to scale the threshold in relation to the mean. The overall mean intensity for control subjects was found to be close to 0.3. Further visual validation showed that the deviation of the mean intensity from 0.3 provided the required threshold correction, hence the inclusion of the term (0.3-μ). However, since the required threshold correction seemed to increase at a faster rate, the term (0.3-μ) was multiplied by the mean to increase the magnitude of the correction factor.

A final observation revealed that in cases with an excessively large or small standard deviation, the mean-corrected threshold proved insufficient. Thus, the standard deviation was factored into the equation to counteract this effect.

Throughout the development process of Eqn. 1, visual validation was carried out by means of a short script file designed to perform segmentation and display all of the images in the dataset in succession. Eqn. 1 was found to produce reasonably accurate segmentation results across the entire dataset.

3.4 Surface Rendering

The thresholding process resulted in 3D binary volumes where voxels within the region of interest were assigned a value of 1 and the remaining were assigned a value of 0. In order to picture the striatal regions, the binary 3D data was transformed into a surface using a 3D interpolation method. Fig. 4 illustrates the 3D striatal region surface render from control and pathological data.

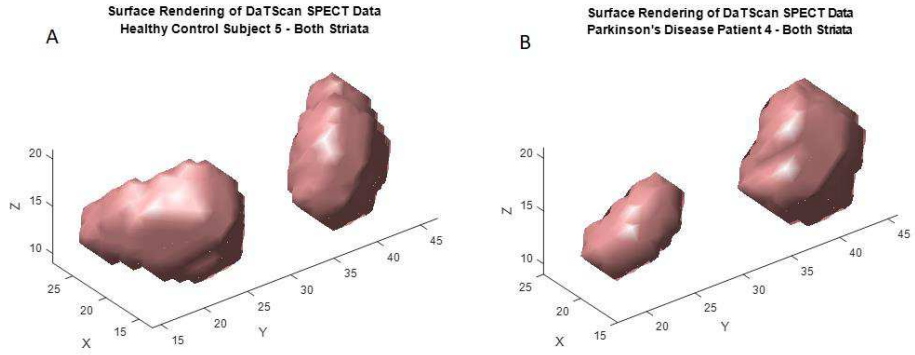

Fig. 4. 3D striatal region surface render from control data (A) and pathological data (B).

3.5 Feature Extraction

Feature extraction involved selecting a set of attributes which can be computed from the data in order to be able to discriminate between healthy and pathological cases. Three types of features were extracted and included in this study: intensity features, shape features and semi-quantitative measures

In a clinical environment, clinicians discriminate between healthy and pathological based on the shape of the uptake region [2] [16]. This warranted the inclusion of shape features into the feature set. Voxel intensity features were included since research suggested that DaT images from Parkinson's disease patients exhibit reduced uptake in the striatal region in comparison with healthy subjects [1]. Due to the initial application for quantifying striatal uptake, semi-quantitative measures were included as features in this study since they are known to differ across the two groups under investigation [16].

The extracted features were the: average and normalised intensity [26], average and normalised intensity gradient, maximum intensity, compactness, diagonal, elongation, extent, major-to-minor-diameter, mid-to-minor diameter [24] [25] [9], total striatal volume and the striatal binding ratio [17]. In each case, the corresponding asymmetry index was also extracted in order to capture the difference between the left and right striatal region in terms of the feature being considered. Asymmetry indexes served as features.

3.6 Feature Selection

Feature selection was used as a dimensionality reduction tool and validation stage to identify a subset of strong features which provide high classifier performance with minimal loss of relevant information. The dataset was split into two equal portions; one of which was used exclusively for feature selection and validation. The remaining portion was used for the final classification. Sequential feature selection was carried out using 10-fold cross-validation. Training and testing of the classifier at feature selection stage

was repeated 10 times with each fold used only once for training. The results were calculated using the misclassification error (MCE) and were averaged over the 10 iterations to provide the final outcome.

Starting with an empty feature set, the feature selection algorithm added a feature to the set and carried out classification using a linear kernel to produce a result before adding another feature and repeating the procedure. The process was repeated iteratively until the MCE was reduced to a minimum. The features forming part of the feature set at the minimum MCE are the selected features. The selected features were the: major-to-minor diameter, elongation, maximum intensity asymmetry index, average intensity asymmetry index, normalised average intensity gradient, striatal volume asymmetry index, extent asymmetry index, striatal binding ratio, compactness asymmetry index and diagonal. The striatal binding ratio is a semi-quantitative measure which was defined as in Eqn. 2, where the striatum is the target region indicated in red in Fig. 5 and the occipital region in a low-uptake reference region indicated in blue in Fig. 5. A detailed explanation of the computation of the mentioned features can be found in [21].

3.7 Classification

Classification was carried out using a SVM with a linear kernel and a feature set made up of the mentioned 10 features. The classifier was trained on 50% of the dataset and tested on the remaining 50%. Classification was repeated 100 times and the accuracy was averaged. Fig. 6 illustrates the pipeline of the system for high-level understanding. The initial three stages of the pipeline were not described in this paper since they were carried out by the PPMI, however they are described in detail in the PPMI manuals [18] [19] [23].

$$SBR = \frac{Mean\ Counts_{Striatum}}{Mean\ Counts_{Occipital\ Region}} - 1 \tag{2}$$

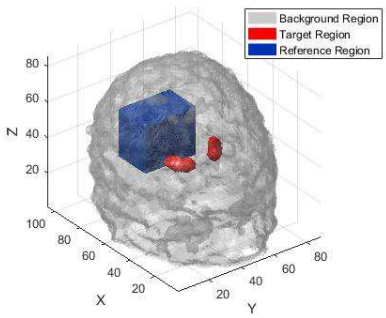

Fig. 5. A 3D illustration indicating the target and reference regions used for SBR calculation.

4 Results

System accuracy was recorded as 97.01%, with 96.63% sensitivity and 97.78% specificity, based on a dataset of 268 test observations. Furthermore, the results indicate that 97% (172) of the pathological cases and 98% (88) of the healthy cases were identified as such whereas 3% (6) of the pathological cases and 2% (2) of the healthy cases were misclassified. Classifier performance was also investigated based in individual features alone in order to identify the discriminative power of each feature. The top three discriminative features were found to be the: striatal binding ratio, diagonal and major-to-minor diameter, as shown in Table 1. Receiver Operating Characteristic (ROC) analyses carried out for each individual feature also confirm that these three features are the ones with the strongest discriminative power. These results also indicate that the striatal uptake regions differ greatly in shape and intensity between the control and pathological groups.

5 Conclusion

The aim of this project was the 3D analysis of DaTScan data and the classification between positive and negative cases of Parkinson's disease. The presented classification system was reported to provide an average accuracy of 97.01% with 96.63% sensitivity and 97.78% specificity.

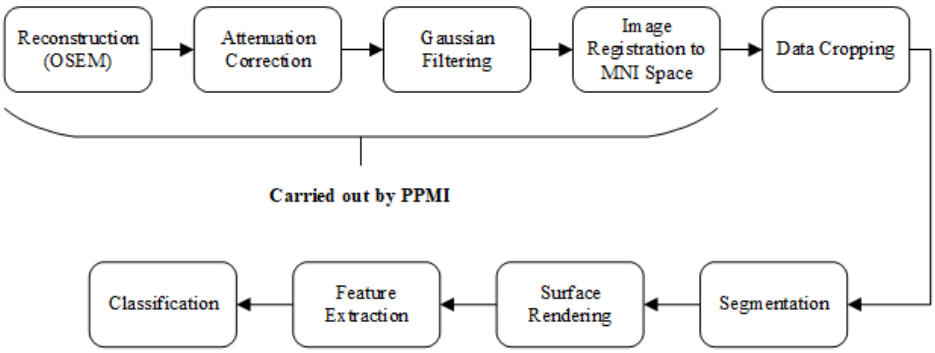

Fig. 6. The system pipeline illustrating the processing steps at a high level.

Table 1. Performance results for the top three discriminative features.

Feature	Accuracy	Sensitivity	Specificity	AUC
Striatal Binding Ratio	84.33%	0.87	0.8	0.773
Diagonal	83.58%	0.83	0.84	0.725
Major-to-Minor Diameter	83.21%	0.78	0.94	0.8111

This compares to the reported 94% sensitivity and 92% specificity for visual interpretation which is the clinical standard [27]. Moreover, it demonstrates an improvement over the use of 2D features which rendered overall system accuracy of 90%, as reported by Prashanth et al. [28].

Three main improvements in system performance have been identified: the elimination of redundancy in the dataset, the use of orthogonal features and the addition of clinical data, such as Hoehn and Yahr ratings.

With the suggested improvements, it is believed that the system would be capable of serving as an aid for Parkinson's disease diagnosis by minimising variability and inaccurate treatment whilst enabling the assessment of disease progression. Further information about this project can be found in [21].

References

1. Simuni T., Rajesh P.: Parkinson's Disease. Oxford University Press, Cary, North Carolina, USA (2009).
2. Seifert K., Wiener J.: The Impact of DaTscan on the Diagnosis and Management of Movement Disorders: A Retrospective Study. American Journal of Neurodegenerative Disease 2(1), 29-34 (2013).
3. Brooks D.: Optimizing Levodopa Therapy for Parkinson's Disease with Levodopa/Carbidopa/Entacapone: Implications from a Clinical and Patient Perspective. Neuropsychiatric Disease and Treatment 4(1), 39-47 (2008).
4. Hughes A., Ben-Shlomo Y., Daniel S., Lees A.: What Features Improve the Accuracy of Clinical Diagnosis in Parkinson's Disease: A Clinicopathologic Study. Journal of Neurology 42(1), 1142-1146 (1992).
5. Hughes A., Daniel S., Lees A.: Improved Accuracy of Clinical Diagnosis of Lewy Body Parkinson's Disease. Journal of Neurology 57(1), 1497-1499 (2001).
6. Marshall V., Reininger C., Marquardt M., Patterson J., Hadley D., Oertel W.: Parkinson's Disease is Overdiagnosed Clinically at Baseline in Diagnostically Uncertain Cases: A 3-year European Multicenter study with Repeat (^{123}I)FP-CIT SPECT. Movement Disorders Journal 24(1), 500-508 (2009).
7. Schwingenschuh P., Ruge D., Edwards M., Terranova C., Katschnig P., Carrillo F.: Distinguishing SWEDD Patients with Asymmetric Resting Tremor from Parkinson's Disease: A Clinical and Electrophysiological Study. Movement Disorders Journal 25(2), 560-569 (2010).
8. DeMaagd G., Philip A.: Parkinson's Disease and Its Management. Pharmacy and Therapeutics 40(11), 747-753 (2015).
9. Prashanth R., Dutta Roy S., Mandal P., Ghosh S.: High-Accuracy Classification of Parkinson's Disease Through Shape Analysis and Surface Fitting in ^{123}I-Ioflupane SPECT Imaging. IEEE Journal of Biomedical and Health Informatics 21(3), 794-802 (2017).
10. Booth T., Nathan M., Waldman A., Quigley A., Schapira A., Buscombe J.: The Role of Functional Dopamine-Transporter SPECT Imaging in Parkinsonian Syndromes. American Journal of Neuroradiology (2014).

11. Cummings J., Henchcliffe C., Schaier S., Simuni T., Waxman A., Kemp P.: The Role of Dopaminergic Imaging in Patients with Symptoms of Dopaminergic System Neurodegeneration. Brain 134(1), 3146-3166 (2011).
12. Valli A., Wiselin Jiji Dr. G.: Parkinson's Disease Diagnosis Using Image Processing Techniques - A Survey. International Journal on Computational Sciences & Applications (IJCSA) 4(6), 57-67 (2014).
13. Sawle G., Playford E., Burn D., Cunnigham V., Brooks D.: Separating Parkinson's Disease from Normality: Discriminant Function Analysis of 18F-DOPA PET Data. Archives of Neurology 51(1), 237-243 (1994).
14. Seibyl J., Marek K., Quinlan D.: Decreased Single-Photon Emission Computed Tomographic ^{123}IB-CIT Striatal Uptake Correlates with Symptom Severity in Parkinson's Disease. Annals of Neurology 38(1), 589-598 (1995).
15. Benamer H., Patterson J., Grosset D.: Accurate differentiation of Parkinsonism and Essential Tremor using Visual Assessment of ^{123}I-FP-CIT SPECT Imaging: The ^{123}I-FP-CIT Study Group. Movement Disorders 15(3), 503-510 (2000).
16. Morton R., Guy M., Clauss R., Hinton P., Marshall C., Clarke E.: Comparison of Different Methods of DaTSCAN Quantification. Nuclear Medicine Communications 26(12), 1139-1146 (2005).
17. European Association of Nuclear Medicine: Brain Imaging - A Technologist's Guide. (2016).
18. Parkinson's Progression Markers Initiative, http://www.ppmi-info.org/, last accessed: June 16,.
19. Marek K.: The Parkinson's Progression Markers Initiative (PPMI) Study Protocol. (2016).
20. Parkinson's Progression Markers Initiative: Imaging Technical Operations Manual. (2014).
21. Azzopardi V.: Analysing Small Brain Structures on DaTScan SPECT Images in Suspected Parkinson's Disease. MSc. Medical Imaging, Centre for Vision Speech and Signal Processing, University of Surrey, 2017, Analysing Small Brain Structures on DaTScan SPECT Images in Suspected Parkinson's Disease, https://www.researchgate.net/project/Analysing-Small-Brain-Structures-on-DaTScan-SPECT-Images-in-Suspected-Parkinsons-Disease, last accessed: 14 April 2018.
22. Marek K, Jennings D, Lasch S, Siderowf A, Tanner C, Simuni T, *et al.*: The Parkinson Progression Marker Initiative (PPMI). Progress in Neurobiology 95(4), 629-635 (2011).
23. Parkinson's Progression Markers Initiative: PPMI Operations Manual. (2015).
24. Lyra M., Striligas J., Gavrilleli M., Lagopati N.: Volume quantification of ^{123}I-DaTSCAN imaging by MatLab for the differentiation and grading of parkinsonism and essential tremor In: INTERNATIONAL CONFERENCE ON SCIENCE AND SOCIAL RESEARCH (CSSR), pp. 163-168. Kuala Lumpur, Malaysia (2010).
25. Prashanth R., Roy S., Ghosh S., Mandal P.: Shape Features as Biomarkers in Early Parkinson's Disease In: 6TH INTERNATIONAL IEEE/EMBS CONFERENCE ON NEURAL ENGINEERING (NER), pp. 517-520. San Diego, California, USA (2013).

26. Segovia F., Illan I., Gorriz J., Ramirez J., Rominger A., Levin J.: Distinguishing Parkinson's Disease from Atypical Parkinsonian Syndromes Using PET Data and A Computer System Based on Support Vector Machines and Bayesian Networks. Frontiers in Computational Neuroscience 9(137), 1-9 (2015).
27. Jakobson S., Linder J., Forsgren L., Riklund K.: Accuracy of Visual Assessment of Dopamine Transporter Imaging in Early Parkinsonism. Movemet Disorders 2(1), (2014).
28. Prashanth R., Dutta Roy S., Mandal P., Ghosh S.: Automatic Classification and Prediction Models for Early Parkinson's Disease Diagnosis from SPECT Imaging. Expert Systems with Applications 41(7), 3333-3342 (2014).