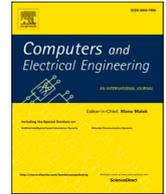




ELSEVIER

Contents lists available at ScienceDirect

## Computers and Electrical Engineering

journal homepage: [www.elsevier.com/locate/compeleceng](http://www.elsevier.com/locate/compeleceng)

# Automated segmentation of substantia nigra and red nucleus using quantitative susceptibility mapping images: Application to Parkinson's disease

Dibash Basukala<sup>a,\*</sup>, Ramakrishnan Mukundan<sup>a</sup>, Anthony Lim<sup>b</sup>, Michael A Hurrell<sup>b</sup>, Ross J Keenan<sup>b,c,d</sup>, John C Dalrymple-Alford<sup>d,e,f,g</sup>, Tim J Anderson<sup>d,f,g,h</sup>, Daniel J Myall<sup>d</sup>, Tracy R Melzer<sup>d,f,g</sup>

<sup>a</sup> Department of Computer Science and Software Engineering, University of Canterbury, Christchurch, New Zealand

<sup>b</sup> Department of Radiology, Christchurch Hospital, Christchurch, New Zealand

<sup>c</sup> Pacific Radiology Group, Christchurch, New Zealand

<sup>d</sup> New Zealand Brain Research Institute, Christchurch, New Zealand

<sup>e</sup> Department of Psychology, University of Canterbury, Christchurch, New Zealand

<sup>f</sup> Brain Research New Zealand Rangahau Roro Aotearoa Centre of Research Excellence, Christchurch, New Zealand

<sup>g</sup> Department of Medicine, University of Otago, Christchurch, New Zealand

<sup>h</sup> Department of Neurology, Christchurch Hospital, Christchurch, New Zealand

## ARTICLE INFO

## Keywords:

Quantitative susceptibility mapping  
Segmentation  
Substantia nigra  
Red nucleus  
Iron deposition  
Parkinson's disease

## ABSTRACT

Accurate segmentation of substantia nigra (SN) and red nucleus (RN) is challenging, yet important for understanding health problems like Parkinson's disease (PD). This paper proposes an algorithm to segment SN and RN from quantitative susceptibility mapping (QSM) MRI and use the results to investigate PD. Algorithm-derived segments (based on level set and watershed transform) are compared to expert manually-derived segmentations in 40 participants. Using Bayesian regression models, we compare QSM values between PD and control groups, and investigate relationships with global cognitive ability and motor severity in PD. The proposed algorithm produces high quality segmentations, validated against expert manual segmentation. We show moderate evidence of increased QSM values in SN in PD relative to controls, with moderate evidence for association between QSM, global cognitive ability, and motor impairment in the SN in PD. We suggest an improved midbrain segmentation algorithm may be useful for monitoring iron-related disease severity in Parkinson's.

## 1. Introduction

Parkinson's disease (PD) is a neurodegenerative disease involving the loss of dopaminergic neurons in the substantia nigra (SN). This cell death affects purposeful movement and eventually results in loss of motor coordination. By the time of diagnosis, approximately 60–80% of dopaminergic neurons within the SN have been lost. In addition to cell death, the SN also experiences increased iron accumulation [1], which may play a role in the pathogenesis of PD. Though its exact role remains unclear, iron accumulation may

This paper was recommended for publication by associate editor Dr. Hui Tian.

\* Corresponding author.

E-mail address: [dibash.basukala@pg.canterbury.ac.nz](mailto:dibash.basukala@pg.canterbury.ac.nz) (D. Basukala).

<https://doi.org/10.1016/j.compeleceng.2021.107091>

Received 2 May 2020; Received in revised form 2 December 2020; Accepted 9 March 2021

Available online 20 March 2021

0045-7906/© 2021 Elsevier Ltd. All rights reserved.

encourage oxidative stress, alpha synuclein aggregation, and the formation of Lewy pathology [2].

Quantitative susceptibility mapping (QSM), an emerging magnetic resonance imaging (MRI) technique, allows high-resolution quantification of magnetic susceptibility, producing images of the small midbrain structures with excellent contrast [3]. Iron is highly paramagnetic and therefore induces a large local magnetic distortion, which is detectable with QSM in vivo. Recent work employing QSM has reported increased susceptibility in PD relative to controls in basal ganglia structures, including the SN and red nucleus (RN), and even cortical regions [4, 5].

The red nucleus is a pair of small gray matter structures which lies adjacent to the substantia nigra in the midbrain. It is also an iron rich structure and is thought to be involved in motor coordination and cognitive functions. Several studies have reported increased iron deposition in the RN of PD patients [6, 7] using QSM, R2\*, and susceptibility-weighted imaging (SWI).

Accurate and precise identification of the structures is key to understanding the processes underlying PD progression in the SN and RN. However, segmentation of these nuclei is challenging because of their small size, unclear boundaries, morphometric variability and similar intensity profiles with the adjacent structures on standard T1- and, to a lesser extent, T2-weighted MRI. Furthermore, tremor is one of the defining characteristics of PD, and therefore motion on all types of brain imaging can confound identification of these, and other, structures. QSM in these midbrain structures (and other subcortical structures) exhibits excellent image contrast, reflecting their intrinsic levels of iron content. Previous work has implemented expert manual tracing of SN [8], but an accurate automated segmentation is preferable to manual segmentation because of the absence of operator bias, the potential for greater consistency, and application to large cohorts in a systematic way. However, there is limited literature available for automated segmentation of these midbrain structures [9–15], potentially due to the relatively recent development of high quality, high contrast imaging capable of visualizing these smaller structures. Currently, atlas-based methods are the most popular method for segmentation of midbrain nuclei. These are performed either in standardized space (warping subjects to normalized space) or in subject space (inverse warping standard regions to subject space) and depend on accurate and robust warping protocols. However, morphometric variability of midbrain structures limits the accuracy of these segmentation methods. Moreover, many methods proposed in the literature require a substantial number of ‘gold standard’ reference images for training, which is computationally expensive and not always feasible. Therefore, work presented in this paper addresses these issues and focused on developing the segmentation method which are independent of atlas registration or label-fusion based methods.

In order to effectively investigate the relationship between iron accumulation and PD, we developed and evaluated a new segmentation algorithm that uses information from QSM images to segment both SN and RN. After comparing algorithm-derived segmentations to ‘gold standard’ segments manually identified by experienced radiologists, we investigated differences in QSM values between PD and controls, and associations with cognitive and motor dysfunction in PD, from both gold standard and algorithm-derived SN and RN segmentations.

The rest of the paper is organized as follows. The materials and methodology used in this work is introduced in Section 2. The results are presented in Section 3, followed by discussion in Section 4. Finally, the paper is summarized in Section 5.

## 2. Materials and methods

A convenience sample of 20 participants meeting the UK Parkinson’s Disease Society’s criteria for idiopathic PD was recruited from volunteers at the Movement Disorders Clinic at the New Zealand Brain Research Institute, Christchurch, New Zealand. Exclusion criteria included atypical Parkinsonian disorders; prior learning disability; previous history of other neurological conditions including moderate-severe head injury, stroke, vascular dementia; and major psychiatric or medical illness in the previous 6 months. Twenty healthy controls matched to the PD group by mean age and sex ratio, were also recruited.

Participants completed neuropsychological testing, clinical assessment, and an MRI scanning session. All participants gave written informed consent, with additional consent from a significant other when appropriate. The study was approved by the regional Ethics Committee of the New Zealand Ministry of Health (No. URB/09/08/037).

### 2.1. Diagnostic criteria and assessment

Motor function was assessed using the Movement Disorders Society Unified Parkinson’s Disease Rating Scale part 3 (UPDRS-III) and cognition by comprehensive neuropsychological testing fulfilling the Movement Disorders Society (MDS) Task Force Level II criteria [16]. Five cognitive domains were examined (executive function; attention, working memory and processing speed; learning and memory; visuospatial/visuoperceptual function; and language; see Supplementary Table 1 for a list of the specific tests). Within each cognitive domain, standardized scores from the constituent neuropsychological tests were averaged to provide individual cognitive domain scores; global cognitive performance for each participant was expressed as an aggregate  $z$  score obtained by averaging all domain scores. Of the 20 PD participants, seven had normal cognition, eight had mild cognitive impairment, and five had dementia. Motor, cognitive, and MRI assessments were performed on medication, with no change to patients’ usual drug regimen.

### 2.2. MRI acquisition

All 40 participants (20 healthy and 20 PD patients) were scanned on a 3T General Electric HDxt scanner (GE Healthcare, Waukesha, USA) with an eight channel head coil. To quantify magnetic susceptibility in the brain, a 3D spoiled gradient recalled echo (SPGR) acquisition with 8 echoes was used to obtain real and imaginary pairs with the following parameters: echo time (TE) = 3.5, 7.3, 11.1, 14.9, 18.7, 22.5, 26.3, 30.2 ms (3.8 ms intervals), repetition time (TR) = 42.8 ms, flip angle = 20 deg, acquisition matrix = 512 × 512

$\times 60$ , field of view (FOV) = 240 mm, slice thickness = 2 mm, voxel size =  $0.47 \times 0.47 \times 2 \text{ mm}^3$ . In addition, we acquired a conventional T1-weighted 3D SPGR acquisition (TE/TR = 2.8/6.6 ms, inversion time (TI) = 400 ms, flip angle = 15 deg, acquisition matrix =  $256 \times 256 \times 170$ , FOV = 250 mm, slice thickness = 1 mm, voxel size =  $0.98 \times 0.98 \times 1.0 \text{ mm}^3$ ), a T2-weighted, and T2 weighted fluid-attenuated inversion recovery images to facilitate a clinical read. All MRI scans were screened by a consultant neuroradiologist (RJK) to exclude significant non-PD cerebral pathology.

### 2.3. QSM reconstruction

Morphology Enabled Dipole Inversion (MEDI) algorithm was used to generate quantitative susceptibility maps from the real/imaginary pairs [17]. The total field was estimated using a nonlinear field map estimation followed by graph-cut-based phase unwrapping [18]. We used the Laplacian-boundary value technique for background field removal [19] and morphology enabled dipole inversion with automatic uniform ventricular cerebrospinal fluid zero referencing (MEDI+0) to produce the quantitative susceptibility maps. A brain mask was extracted from the averaged magnitude image using FSL's (v6.0.3) Brain Extraction Tool (BET).

CAT12 (r1278, <http://www.neuro.uni-jena.de/cat/>), a toolbox of SPM12 (v7219, <http://www.fil.ion.ucl.ac.uk/spm/>), running in MATLAB 9.3.0 (R2017b), was used to process T1-weighted structural images. Each individual's T1-weighted structural image was first re-oriented (using SPM12's coregistration routine) to be in rough alignment with Montreal Neurological Institute (MNI) space; the average magnitude image was then coregistered to the T1-weighted structural image, along with each individual QSM map, to ensure a relatively consistent orientation in subject space.

T1-weighted images were bias corrected, spatially normalized via DARTEL (using the MNI-registered template provided within CAT12), modulated to compensate for the effect of spatial normalization, and classified into gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF), all within the same generative model. QSM images were warped into MNI space using the T1-derived deformation fields and averaged to create a study-specific, average QSM image.

### 2.4. Substantia nigra seeding

On the normalized, study-specific average QSM image, a single voxel was manually placed within the SN on each of six slices, creating left and right SN seed images in MNI space (MNI coordinates are given in Supplementary Table 2). These MNI-space seeds were then inverse-warped into QSM subject space using the inverse deformation fields produced during the structural T1-weighted image processing. This produced seeds within the left and right SN in each individual; these seeds were used to initiate the SN and RN segmentations.

### 2.5. Segmentation method

Full details of the segmentation method are described in the following conference paper [20]. Briefly, our method consisted of the following steps:

#### A. Contrast enhancement

The first step in the segmentation process was contrast enhancement. The proposed contrast enhancement technique was based on the local and global gray level information of the input image. The global gray level information was computed by calculating the global mean  $g_{mean}$  of an input image  $A(x, y)$  from its entire sample as follows:

$$g_{mean} = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N A(x, y) \quad (1)$$

where  $M, N$  denote the image size (width, height respectively) in pixels.

The local gray level information was estimated by computing the mean of each non-overlapping  $3 \times 3$  sub-image using an average filter (local mean). Let the sub-image of specified size be centered on  $(x, y)$ . The contrast enhancement was performed on each sub-image as follows:

$$I(x+i, y+j) = A(x+i, y+j) + [I_{mean}(x, y) - g_{mean}] \quad (2)$$

where  $i, j = -1, 0, 1$ ,  $A(x+i, y+j)$  is the input image,  $I_{mean}(x, y)$  is the local mean at  $(x, y)$ , and  $I(x+i, y+j)$  is the contrast enhanced sub-image. The overall contrast enhanced image  $I(x, y)$  was computed by repeating the process in (2) for all the sub-images.

#### B. Level set method

Li et al. [21] proposed a model by embedding the local image information, which helps to segment images with intensity inhomogeneities, common in medical images. The energy functional for this model is given by,

$$F(\phi, f_1, f_2) = \sum_{i=1}^2 \lambda_i \int \left( \int K_\sigma(x-y) |I(y) - f_i(x)|^2 M_i^e(\phi(y)) dy \right) dx \quad (3)$$

$$+ \nu \int |\nabla H_e(\phi(x))| dx + \mu \int \frac{1}{2} (|\nabla \phi(x)| - 1)^2 dx$$

where  $f_1(x)$  and  $f_2(x)$  approximate the local intensities in the regions outside and inside the contour respectively.  $K_\sigma$  is a Gaussian kernel with a scale parameter  $\sigma > 0$ .  $H_\epsilon$  is a smooth Heaviside function,  $M_1^\epsilon(\phi) = H_\epsilon(\phi)$  and  $M_2^\epsilon(\phi) = 1 - H_\epsilon(\phi)$ ,  $v \geq 0$  and  $\mu$  is a positive constant. The minimization of energy functional  $F(\phi, f_1, f_2)$  with respect to  $\phi$  was obtained by the gradient descent method,

$$\frac{\partial \phi}{\partial t} = -\delta_\epsilon(\phi)(\lambda_1 e_1 - \lambda_2 e_2) + v \delta_\epsilon(\phi) \operatorname{div} \left( \frac{\nabla \phi}{|\nabla \phi|} \right) + \mu \left( \nabla^2 \phi - \operatorname{div} \left( \frac{\nabla \phi}{|\nabla \phi|} \right) \right) \tag{4}$$

where  $\delta_\epsilon$  is the smoothed Dirac delta function.

**C. Wavelet transform**

In our work, we used a dual-tree complex wavelet transform (DT-CWT) for the purpose of smoothing.

*Dual-tree complex Wavelet transform*

The one-dimensional (1-D) DT-CWT [22] decomposes an input signal  $f(x)$  in terms of a complex shifted and dilated mother wavelet  $\psi(x)$  and scaling function  $\phi(x)$ . Mathematically,

$$f(x) = \sum_{l \in Z} B_{j_0, l} \phi_{j_0, l}(x) + \sum_{j \geq j_0} \sum_{l \in Z} w_{j, l} \psi_{j, l}(x) \tag{5}$$

where  $Z$  is the set of natural numbers,  $j_0$  represents the number of decomposition levels,  $j$  and  $l$  refer to the index of shifts and dilations, respectively.  $B_{j_0, l}$  and  $\phi_{j_0, l}(x)$  are the scaling coefficient and scaling function, respectively.  $w_{j, l}$  represents the complex wavelet coefficient, while  $\psi_{j, l}(x)$  is the complex wavelet function.

Similarly, the two-dimensional (2-D) DT-CWT decomposes an image  $f(x, y)$  through a sequence of dilations and translations of a complex scaling function and six complex wavelet functions. Mathematically,

$$f(x, y) = \sum_{l \in Z^2} B_{j_0, l} \phi_{j_0, l}(x, y) + \sum_{\theta \in \beta} \sum_{j \geq j_0} \sum_{l \in Z^2} w_{j, l}^\theta \psi_{j, l}^\theta(x, y) \tag{6}$$

where  $\theta \in \beta = \{ \pm 15^\circ, \pm 45^\circ, \pm 75^\circ \}$  provides the directionality of the complex wavelet function.

**D. Watershed transformation**

The watershed transform is a popular segmentation method that has been widely used in medical image segmentation. It is a simple, fast, and intuitive method that produces a complete division of the image into distinct regions, even when the contrast is poor, thus avoiding the need for any kind of post-processing work, such as contour joining.

Distance transform was computed before applying the watershed transform. The watershed transformation subsequently performed on the distance transformed image results in over-segmentation—the main drawback of this algorithm—because of the presence of several local minima. Therefore, an extended-minima transform was used to modify the distance transformed image for accurate and effective segmentation.

*Extended-minima transform*

The extended minima transform was defined as the regional minima of the corresponding  $H$ -minima transformation. Mathematically,

$$EMIN_h(f) = RMIN[HMIN_h(f)] \tag{7}$$

where,

$$HMIN_h(f) = R_f^\epsilon(f + h) \tag{8}$$

is the  $H$ -minima transformation and is obtained by performing the reconstruction by erosion of  $f$  with respect to  $f + h$ .

**2.6. Implementation**

The bounding box for SN and RN segmentation was obtained by substantia nigra seeding. Contrast enhancement was performed on the values within the bounding box using (2) for all the sub-images. The contrast enhanced image helped in directing the evolution of the level set function towards the desired direction.

The boundaries of the segmentations were obtained with the level set method while DT-CWT smoothed the jagged output and resolved over-segmentation resulting from the level set method. Watershed transformation was performed on the modified distance transformed image to separate and segment SN and RN effectively. The implementation detail of the algorithm can be found in [20].

**2.7. Qualitative assessment of segmentations**

The SN and RN were segmented using three different methods: (1) our proposed method, (2) a level set method based on minimization of region scalable fitting (RSF) energy for segmentation of images with intensity inhomogeneities proposed by Li et al. [21], and (3) a different level set method proposed by Li et al. [23] that is based on the local intensity clustering and is capable of handling images with intensity inhomogeneities. We selected these two level set methods as comparisons as they are established methods for segmentation of smaller structures in medical imaging.

Two experienced radiologists separately performed subjective evaluation of the segmentations produced by the three methods, described above. Subjective evaluation was performed blinded to method, group, and demographic information. A score ranging from 1 to 5 (very poor, poor, fair, good, very good) was used for the subjective assessment of the segmentations.

## 2.8. Quantitative assessment: expert ('gold standard') manual segmentations

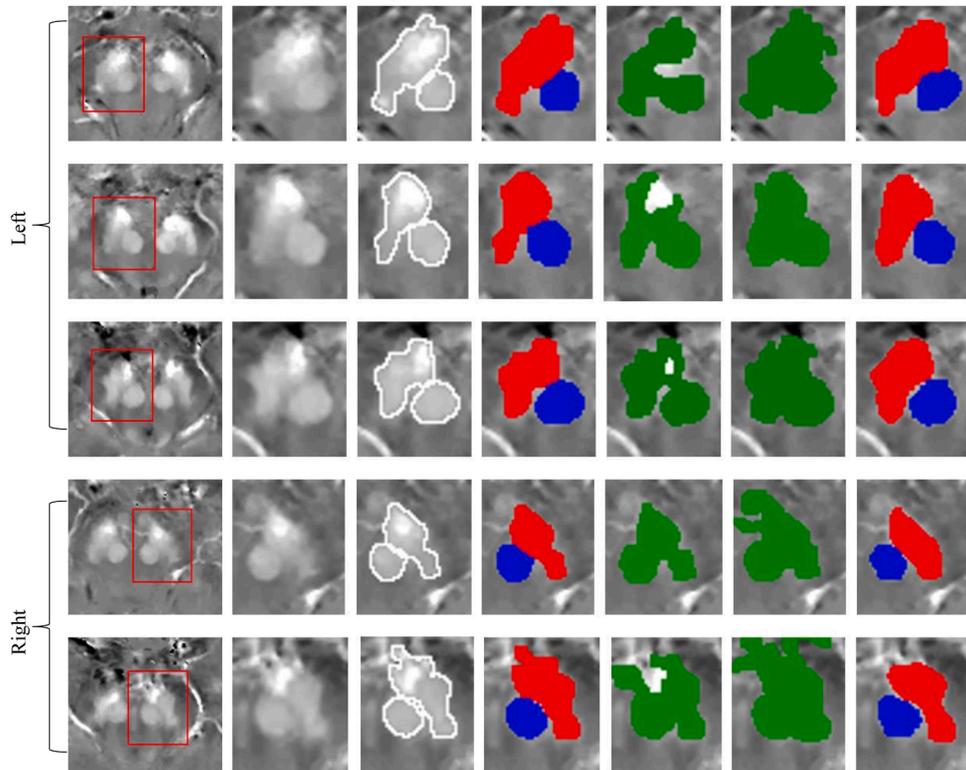
An experienced radiologist (AL) performed the manual segmentations of SN and RN on all 40 subjects. A second experienced radiologist (MH) verified the manual segmentations drawn by the first radiologist; in the case of disagreement, the manual segmentation was adjusted until both radiologists were in agreement. Radiologists were blinded to all clinical and demographic characteristics. These manual segmentations were considered as the gold standard for further analyses.

Dice score [24], a commonly used similarity metric to measure the spatial overlap between two segmentations, was used to assess segmentation accuracy of our proposed method; specifically, manual segmentations performed by the radiologists were compared with the results generated by our automated segmentation method. The value of Dice score ranges between 0 (no overlap) and 1 (complete overlap).

Median QSM values were extracted from gold standard SN and RN segments, as well as from segments produced with our proposed method (right and left were extracted independently). Pearson correlation coefficients were used to assess the correlation between QSM values extracted from the two methods; reliability was measured using the intraclass correlation coefficient (ICC) type (A,1) and (C,1) [25].

## 2.9. Parkinson's disease-related differences in susceptibility

Median QSM values from both 'gold standard' and our proposed method SN and RN ROIs were compared across group, global cognitive ability, and motor impairment using Bayesian regression models. Bayesian models were fitted using the "brms" (v2.9.0) package in R (v3.6.3). In each model, four chains with 5000 iterations each were used to generate the posterior sample. Given previous



**Fig. 1.** Results of the proposed segmentation method and comparison with different methods. Column 1: Bounding box obtained by substantia nigra seeding. The box is represented in red. Column 2: Area within the bounding box to initiate the segmentation. Column 3: Segmentation result of our proposed method. The result of automated segmentation is shown as a white contour. Column 4: labeled mask of the proposed segmentation method is shown in red (SN) and blue (RN). Column 5: Result of the level set method based on minimization of region-scalable fitting energy. Column 6: Result of the level set method in the presence of intensity inhomogeneities with application to MRI. SN and RN both are labeled as green since level set method cannot separate them. Column 7: Manually labeled structures of SN (red) and RN (blue). SN = Substantia nigra, RN = Red nucleus. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

research, we examined hypotheses of increased QSM in PD relative to controls, a negative association with global cognitive ability, and a positive association with motor impairment. Parameter estimates, 95% uncertainty intervals, and probabilities of hypotheses being true are presented.

- **Group differences:** We first tested for evidence of varying susceptibility (QSM value) between PD and healthy control groups. For each segmentation method (gold standard/proposed method), we modeled QSM as a function of age, sex, group (PD/control), side (left or right), ROI (SN/RN), and the interactions age-by-ROI (allowing the relationship between age and QSM to vary by ROI), group-by-ROI (allowing the difference in QSM between PD and controls to vary between ROI), and group-by-side (allowing the PD effect to vary between left and right side), with an intercept varying by subject, side, and ROI.
- **Cognitive ability:** In PD participants only, we modeled QSM as a function of age, sex, side (left or right), ROI (SN/RN), global cognitive ability (aggregate cognitive  $z$  score), and the interactions age-by-ROI, cognitive score-by-ROI, and cognitive score-by-side, with an intercept varying by subject, side, and ROI.
- **Motor impairment:** In PD participants only, we modeled QSM as a function of age, sex, side (left or right), ROI (SN/RN), UPDRS III, and the interactions age-by-ROI, UPDRS III-by-ROI, and UPDRS III-by-side, with an intercept varying by subject, side, and ROI.

### 3. Results

We used the following parameters for the level set method in our experiments:  $\lambda_1 = 1.0$ ,  $\lambda_2 = 2.0$ ,  $\mu = 1$ ,  $\nu = 0.003 \times 255 \times 255$  and  $\sigma = 10.0$ . The level set parameters were defined after evaluation with several variations, especially the scale parameter  $\sigma$ . A smaller  $\sigma$  could be used in the experiments, but it would require large number of iterations for accurate computation.

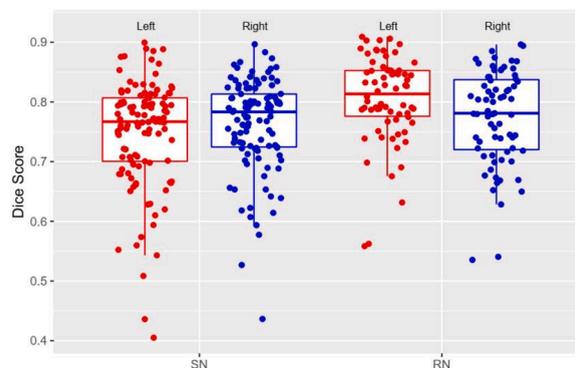
#### 3.1. Algorithm performance

The proposed methodology produced good results. Segmentations using different methods are displayed in Fig. 1. The proposed segmentation method received the highest subjective score from both radiologists. The alternative level set methods did not segment or separate SN and RN effectively. In the quantitative analysis, our proposed method produced good Dice scores (Fig. 2). The Dice score (mean (standard deviation)) of the left SN and right SN between the gold standard segmentation and the proposed method were 0.75 (0.09) and 0.76 (0.07), respectively; left RN and right RN were 0.81 (0.07) and 0.77 (0.07), respectively. Fig. 3 presents the relationship between gold standard and proposed method QSM values in both SN and RN. Slope ( $m$ ) and Pearson correlation coefficients ( $r$ ), with 95% confidence intervals, were as follows: left SN:  $m = 0.94$  [0.82, 1.06],  $r = 0.93$  [0.87, 0.96], right SN:  $m = 0.92$  [0.76, 1.07],  $r = 0.89$  [0.80, 0.94], left RN:  $m = 0.92$  [0.83, 1.00],  $r = 0.96$  [0.93, 0.98], right RN:  $m = 0.98$  [0.87, 1.09],  $r = 0.95$  [0.90, 0.97]. ICC showed moderate-good agreement and excellent consistency (Table 1).

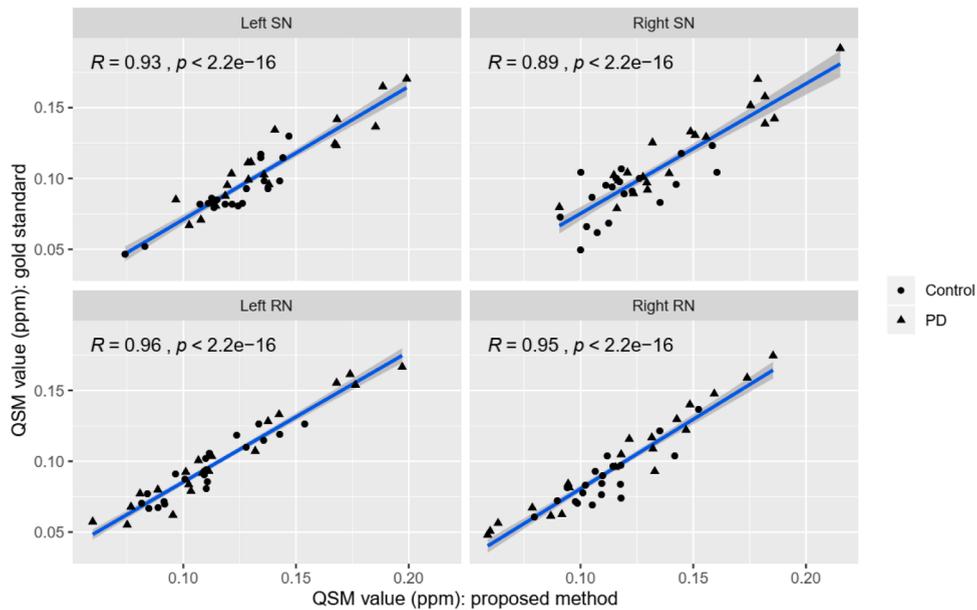
#### 3.2. Application to Parkinson's disease

Demographic and clinical data are summarized in Table 2.

A. Group differences: Table 3 and Fig. 4 report the results from comparisons between healthy control and PD groups in SN and RN for both the gold standard segmentations and those from the proposed method. We observed moderate evidence of increased QSM values in the SN in the PD group relative to controls (probability  $P = 98\%$  (proposed) and  $P = 99\%$  (manual segmentation) of the PD group having higher QSM values than the control group), with weak evidence within the RN ( $P = 75\%$  (proposed) and  $P = 84\%$  (manual). Furthermore, we found moderate evidence for an association with age in the SN (probability  $P = 92\%$  of a positive association, proposed method), but no evidence in the RN ( $P = 29\%$ ), and no evidence of a difference between anatomic side (probability  $P = 48\%$  probability of right side having greater QSM values).



**Fig. 2.** Boxplot of Dice scores of the proposed segmentation method for the left and right sides of SN and RN. Dots are Dice scores for all the slices of images used in the experiment from all 40 subjects. SN = Substantia nigra, RN = Red nucleus. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 3.** Relationship between gold standard QSM values and proposed method QSM values. SN = Substantia nigra, RN = Red nucleus, R = Pearson correlation coefficient,  $p = p$  value. The line of best fit is displayed in blue, with 95% confidence intervals shaded in gray. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Table 1**  
Intraclass correlation coefficient for agreement and consistency.

	ICC(A,1): Left	ICC(A,1): Right	ICC(C,1): Left	ICC(C,1): Right
SN	0.56 [−0.04, 0.86]	0.62 [−0.08, 0.88]	0.93 [0.87, 0.96]	0.89 [0.80, 0.94]
RN	0.84 [−0.04, 0.96]	0.77 [−0.06, 0.94]	0.96 [0.93, 0.98]	0.94 [0.90, 0.97]

Values are ICC [95% CI]. ICC(A,1) = Intraclass correlation coefficient for agreement, ICC(C,1) = Intraclass correlation coefficient for consistency. CI: Confidence interval.

**Table 2**  
Demographic and clinical information.

Measure	Control (n = 20)	Parkinson’s disease (n = 20)
Sex (M:F)	14:6	14:6
Age	72 (8)	74 (6)
Global cognitive ability	0.8 (0.4)	−0.8 (1.0)
MoCA	27 (2)	22 (4)
UPDRS-III	–	41 (12)
Duration of diagnosis (years)	–	8 (5)

Data are mean (SD: standard deviation). Sex is reported as a ratio. M: Male; F: Female. MoCA: Montreal Cognitive Assessment; UPDRS-III: Unified Parkinson’s Disease Rating Scale part 3. Global cognitive ability was calculated as an aggregate of standardized scores across five cognitive domains.

**B. Cognitive impairment, motor impairment, and QSM:** Table 4, Figs. 5 and 6 report the relationship between QSM, global cognitive ability, and motor impairment. We observed moderate evidence of a negative relationship between QSM and global cognitive z score in SN (94% (proposed) and 87% (manual segmentation) probability of a negative association between QSM and global cognitive ability).

**Table 3**  
QSM differences between Parkinson’s disease and control participants.

Region	Gold standard		Proposed method	
	SN	RN	SN	RN
Control group (ppb)	85 [68, 101]	84 [67, 101]	115 [98, 132]	104 [86, 121]
PD vs controls (ppb)	20 [6.1, 34]	8.7 [−5.7, 23]	18 [3.7, 33]	5.9 [−8.9, 20]
Posterior probability of PD > controls	99%	84%	98%	75%

ppb = parts per billion. Posterior probability refers to the probability of PD being greater than controls.

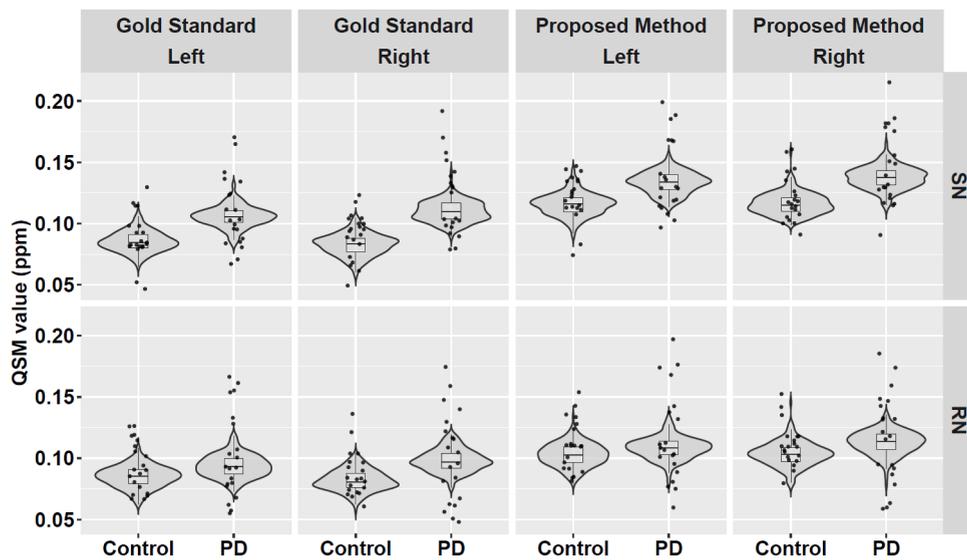


Fig. 4. QSM values of left and right sides of SN and RN between healthy control and PD. Dots are raw QSM values. Violin plots in gray and boxplots represent the distribution of QSM values adjusted for covariates from 200 draws from the posterior probability distribution. SN = Substantia nigra, RN = Red nucleus, PD = Parkinson’s disease. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

There was no evidence of association within RN (66% (proposed) and 61% (manual segmentation)) and no evidence of a difference in association between left and right ( $-0.11$  ppb/cognitive  $z$  point [ $-15, 15$ ]). There was moderate evidence for a positive association with motor impairment in SN (87% (proposed) and 95% (manual segmentation)) and within RN (96% (proposed) and 98% (manual segmentation)).

#### 4. Discussion

In this study, we proposed an effective automated method for SN and RN segmentation by combining contrast enhancement, level set method, DT-CWT and a watershed transform. Segmentations from our proposed method were compared to radiologist-drawn ‘gold standard’ segmentations. The extracted QSM values were then tested in a cohort of 20 healthy and 20 PD participants; consistent results between the two methods suggests our proposed segmentation algorithm may be used in the future with confidence.

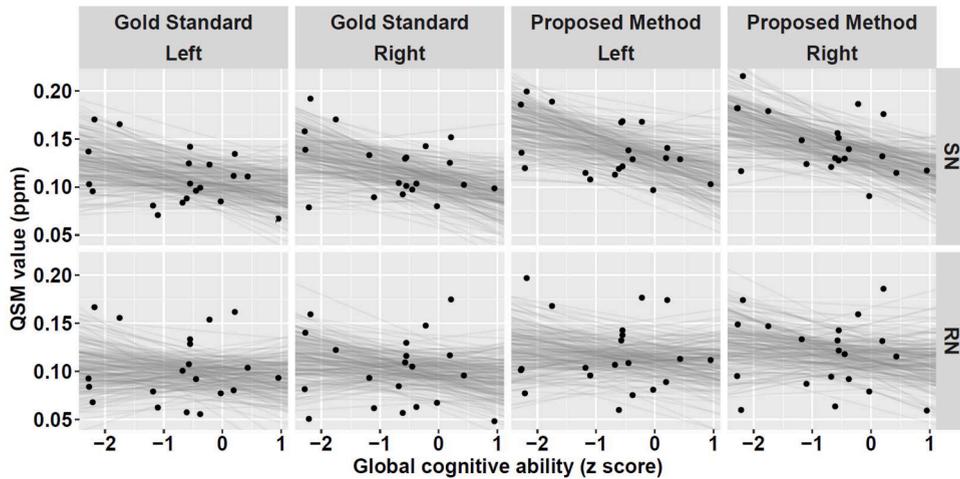
The proposed automated segmentation method of combining various techniques progressed in the following manner. The contrast enhancement technique, which is based on the local and global mean, greatly improved the contrast of the region within the bounding box, which in turn helped in enhancing the contrast of the ROI. Specifically, contrast enhancement helps to increase the accuracy and effectiveness of segmentation, as well as directing the evolution of the level set function towards the desired direction. In our work, the proposed level set method used local image intensities rather than global intensities, which made it capable of handling images with intensity inhomogeneities. Moreover, the level set method also avoided the need for time consuming and expensive re-initialization processes. However, the naive level set method resulted in jagged output, over-segmentation, and did not separate SN from RN in some slices. Therefore, DT-CWT was used while evolving the level set function to address these issues, producing smooth output and resolving over-segmentation. However, mid-brain nuclei were still not effectively separated with this approach. Hence, a watershed transform based on the extended-minima transform was introduced to effectively separate and segment SN and RN. The combination of the above techniques resulted in the best segmentation that adequately matched expert manual segmentation.

The level set [21, 23], a classical and effective segmentation method based on the change of gray level, was employed for comparative analysis in this study because it is frequently used in medical imaging. However, the level set method was unable to separate and segment SN and RN effectively and efficiently, specifically due to the similarity in intensity profile with the adjacent

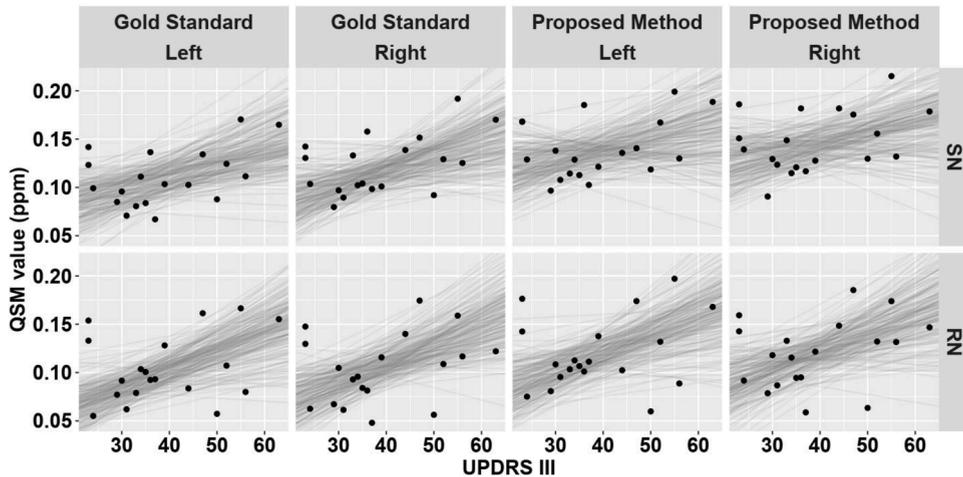
Table 4  
Associations between QSM, global cognitive ability, and motor impairment.

Region	Gold standard		Proposed method	
	SN	RN	SN	RN
Global cognitive ability (ppb/point)	-10 [-25, 5]	-2 [-17, 12]	-14 [-29, 0.8]	-4 [-18, 12]
Posterior probability of a negative association	87%	61%	94%	66%
UPDRS-III (ppb/point)	13 [-0.2, 27]	18 [4, 32]	10 [-5, 25]	17 [2, 31]
Posterior probability of a positive association	95%	98%	87%	96%

The reported value [95% uncertainty interval] is the slope of the association: parts per billion per one point in global cognitive  $z$  score/UPDRS-III, accounting for age, sex, and anatomic side. UPDRS-III = Unified Parkinson’s disease rating scale part 3.



**Fig. 5.** Correlations between QSM values of left and right sides of SN and RN with global cognitive ability (aggregate cognitive  $z$  score) in Parkinson's disease. Black points represent raw values. Lines of best fit, adjusted for covariates, from 200 draws of the posterior probability distribution are displayed in gray. SN = Substantia nigra, RN = Red nucleus, ppm = parts per million. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 6.** Correlations between QSM values of left and right sides of SN and RN with motor impairment (UPDRS-III) in Parkinson's disease. Black points represent raw values. Lines of best fit, adjusted for covariates, from 200 draws of the posterior probability distribution are displayed in gray. SN = Substantia nigra, RN = Red nucleus, UPDRS-III: Unified Parkinson's Disease Rating Scale part 3, ppm = parts per million. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

areas, as well as the blurry boundary that exists between the smaller nuclei. We also observed that the level set method either underestimated or overestimated the SN and RN. Moreover, the level set method resulted in over-segmentation of the midbrain nuclei in the form of holes, which did not make anatomical sense, according to expert clinicians' perspective.

Quantitative assessment was performed using Dice score, the most commonly used similarity metric, to measure the fraction of spatial overlap between binary regions. The Dice score obtained in this study is comparable to acceptable levels in the current literature on SN and RN [9, 12, 15]. Such comparisons must be made with care, however, as image acquisition parameters, quality of manual segmentations, and specific sample heterogeneity may all influence this metric. For example, the partial volume effect and the absolute faint contrast of the midbrain nuclei present in some slices of few subjects limit segmentation accuracy and may reduce the overall Dice score. The Dice score for the level set methods was not calculated as these comparison methods were not able to separate the smaller midbrain structures. While we have taken the expert manual segmentation as the 'gold standard', there is a certain amount of uncertainty associated with these segments, as with any manual segmentation. However, we attempted to minimize this uncertainty by requiring agreement between two experts which resulted in good reliability.

In the current work, we applied our novel segmentation method to a cohort of PD patients. Quantitative susceptibility mapping allows in vivo investigation of iron accumulation in brain tissue, and recent work suggests increased iron deposition in SN in PD [26,

27]. Our findings, of moderate evidence of increased QSM values in PD in the SN and in the RN, are consistent with the expected increase in iron accumulation in PD patients. QSM values with SN and RN exhibited moderate evidence of an association with motor severity in PD. This further suggests that iron accumulation may continue as the disease progresses, providing an imaging marker that tracks motor impairment. While pathology within the SN, and RN to a lesser extent, is generally associated with early disease and the emergence of motor impairments in PD, we also showed a moderate association between QSM values within the SN and global cognitive ability. A widespread pattern of abnormal iron deposition in both cortical and subcortical regions may also underlie cognitive impairments [4, 5]. Interestingly, these two recent studies (identifying SN manually or atlas-based) did not identify correlations with a global cognitive score (the Montreal Cognitive assessment) or the UPDRS-III. The combination of improved segmentation of the SN using the proposed method along with more comprehensive neuropsychological assessments may help explain our moderate associations.

Our study has some limitations. First, the proposed algorithm is primarily based on the intensity and texture of the image. Therefore, the performance of the algorithm drops when the contrast is poor. Thus, better imaging modalities providing adequate contrast of the midbrain nuclei could be used for improved performance and segmentation accuracy. Second, the sample size used in the study for both groups was relatively small, which may not capture the full range of QSM and SN variability during PD progression. Third, we only considered the entire SN, aggregating pars compacta, pars reticulata, and dorsal and ventral sub-regions into the full SN. While dopaminergic cell death is prominent within the pars compacta in PD, studies investigating iron accumulation have reported increase in both the pars compacta and pars reticulata [28]. A recent study [29], assessed iron separately in the ventral and dorsal levels of the SN in PD and specifically found significant iron increment in the ventral SN of the PD patients relative to healthy controls. While the specificity of iron accumulation in specific sub-regions of the SN is of interest and should be explored in the future, we focused on segmentation of the whole SN given its small size, relatively large imaging voxels, and lack of obvious contrast within the QSM images. Thus, we are unable to comment on any differences across the differing SN sub-regions, however future extension of the current segmentation algorithm should allow identification of SN sub-regions.

Having validated the proposed SN and RN segmentation method, we now have the opportunity to apply this technique to a large longitudinal cohort of PD patients and controls, to allow a multi-modal examination including QSM,  $R2^*$ , volume, diffusion tensor imaging (DTI) parameters and cerebral blood flow (CBF) of these two key midbrain nuclei across groups and in the evolution of cognitive and motor decline in PD. Additionally, this segmentation method may be applied to other disease groups, such as those with Alzheimer's disease (AD), given the recent findings of increased iron accumulation in AD [30].

## 5. Conclusion

We presented a new automated method for the segmentation of SN and RN from QSM images by combining contrast enhancement technique, level set method, DT-CWT and watershed transform. The proposed method produced high quality segmentations of SN and RN, validated against expert manual segmentations, and facilitated in vivo assessment of midbrain nuclei iron content. Specifically, we showed increased QSM values in SN in PD relative to controls, and moderate evidence of association between SN and RN QSM values with motor impairment; these results were comparable between both our proposed segmentation and expert manual segmentation. Work presented here suggests that an improved midbrain segmentation algorithm may be useful for monitoring iron-related disease severity in Parkinson's disease.

## Author statement

Conceptualization: Dibash Basukala, Ramakrishnan Mukundan, Tracy R Melzer.

Methodology: Dibash Basukala, Ramakrishnan Mukundan, Anthony Lim, Michael A Hurrell, Daniel J Myall, Tracy R Melzer.

Validation: Dibash Basukala, Anthony Lim, Michael A Hurrell, Ross J Keenan.

Formal analysis: Dibash Basukala, Daniel J Myall, Tracy R Melzer.

Resources: Ross J Keenan, John C Dalrymple-Alford, Tim J Anderson, Tracy R Melzer.

Writing - Original Draft: Dibash Basukala, Daniel J Myall, Tracy R Melzer.

Writing - Review & Editing: Dibash Basukala, Ramakrishnan Mukundan, Anthony Lim, Michael A Hurrell, Ross J Keenan, John C Dalrymple-Alford, Tim J Anderson, Daniel J Myall, Tracy R Melzer.

Supervision: Ramakrishnan Mukundan, Tracy R Melzer

Funding acquisition: John C Dalrymple-Alford, Tim J Anderson, Tracy R Melzer.

## Declaration of Competing Interest

None.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.compeleceng.2021.107091](https://doi.org/10.1016/j.compeleceng.2021.107091).

## References

- [1] Ward RJ, Zucca FA, Duyn JH, Crichton RR, Zecca L. The role of iron in brain ageing and neurodegenerative disorders. *Lancet Neurol* 2014;13(10):1045–60.
- [2] Bjorklund G, Hofer T, Nurchi VM, Aaseth J. Iron and other metals in the pathogenesis of Parkinson's disease: toxic effects and possible detoxification. *J Inorg Biochem* 2019;199:110717.
- [3] de Rochefort L, Liu T, Kressler B, Liu J, Spincemaille P, Lebon V, Wu J, Wang Y. Quantitative susceptibility map reconstruction from MR phase data using Bayesian regularization: validation and application to brain imaging. *Magn Reson Med* 2010;63(1):194–206.
- [4] Thomas GEC, Leyland LA, Schrag A-E, Lees AJ, Acosta-Cabrero J, Weil RS. Brain iron deposition is linked with cognitive severity in Parkinson's disease. *J Neurol Neurosurg Psychiatry* 2020;vol. 91(no. 4):418.
- [5] Uchida Y, Kan H, Sakurai K, Arai N, Kato D, Kawashima S, Ueki Y, Matsukawa N. Voxel-based quantitative susceptibility mapping in Parkinson's disease with mild cognitive impairment. *Mov Disord: Off J Mov Disord Soc* 2019;34(8):1164–73.
- [6] Chen Q, Chen Y, Zhang Y, Wang F, Yu H, Zhang C, Jiang Z, Luo W. Iron deposition in Parkinson's disease by quantitative susceptibility mapping. *BMC Neurosci* 2019;20(1):23. 2019/05/22.
- [7] Wang JY, Zhuang QQ, Zhu LB, Zhu H, Li T, Li R, Chen SF, Huang CP, Zhang X, Zhu JH. Meta-analysis of brain iron levels of Parkinson's disease patients determined by postmortem and MRI measurements. *Sci Rep* 2016;6:36669.
- [8] Acosta-Cabrero J, Cardenas-Blanco A, Betts MJ, Butryn M, Valdes-Herrera JP, Galazky I, Nestor PJ. The whole-brain pattern of magnetic susceptibility perturbations in Parkinson's disease. *Brain* 2017;140(1):118–31.
- [9] Guo T, Song Y, Li J, Fan M, Yan X, He A, Huang D, Shen C, Zhang G, Yang G. Seed point discontinuity-based segmentation method for the substantia nigra and the red nucleus in quantitative susceptibility maps. *J Magn Reson Imaging* 2018;48(4):1112–9.
- [10] Haegelen C, Coupe P, Fonov V, Guizard N, Jannin P, Morandi X, Collins DL. Automated segmentation of basal ganglia and deep brain structures in MRI of Parkinson's disease. *Int J Comput Assist Radiol Surg* 2013;8(1):99–110.
- [11] Kim J, Lenglet C, Duchin Y, Sapiro G, Harel N. Semiautomatic segmentation of brain subcortical structures from high-field MRI. *IEEE J Biomed Health Inform* 2014;18(5):1678–95.
- [12] Visser E, Keuken MC, Forstmann BU, Jenkinson M. Automated segmentation of the substantia nigra, subthalamic nucleus and red nucleus in 7 T data at young and old age. *Neuroimage* 2016;139:324–36.
- [13] Xiao Y, Bailey L, Chakravarty MM, Beriault S, Sadikot AF, Pike GB, Collins DL. Atlas-based segmentation of the subthalamic nucleus, red nucleus, and substantia nigra for deep brain stimulation by incorporating multiple MRI contrasts. *Inf Process Comput-Assist Interv (IPCAI)* 2012:135–45. June.
- [14] Xiao Y, Jannin P, D'Albis T, Guizard N, Haegelen C, Lalys F, Verin M, Collins DL. Investigation of morphometric variability of subthalamic nucleus, red nucleus, and substantia nigra in advanced Parkinson's disease patients using automatic segmentation and PCA-based analysis. *Hum Brain Mapp* 2014;35(9):4330–44.
- [15] Garzon B, Sitnikov R, Bäckman L, Kalpouzos G. Automated segmentation of midbrain structures with high iron content. *Neuroimage* 2018;170:199–209.
- [16] Litvan I, Goldman JG, Tröster AI, Schmand BA, Weintraub D, Petersen RC, Mollenhauer B, Adler CH, Marder K, Williams-Gray CH, Aarsland D, Kulisevsky J, Rodriguez-Oroz MC, Burn DJ, Barker RA, Emre M. Diagnostic criteria for mild cognitive impairment in Parkinson's disease: movement Disorder Society Task Force guidelines. *Mov Disord* 2012;27(3):349–56.
- [17] Liu Z, Spincemaille P, Yao YH, Zhang Y, Wang Y. MEDI+0: morphology enabled dipole inversion with automatic uniform cerebrospinal fluid zero reference for quantitative susceptibility mapping. *Magn Reson Med* 2018;79(5):2795–803.
- [18] Dong J, Liu T, Chen F, Zhou D, Dimov A, Raj A, Cheng Q, Spincemaille P, Wang Y. Simultaneous phase unwrapping and removal of chemical shift (SPURS) using graph cuts: application in quantitative susceptibility mapping. *IEEE Trans Med Imaging* 2015;34(2):531–40.
- [19] Zhou D, Liu T, Spincemaille P, Wang Y. Background field removal by solving the Laplacian boundary value problem. *NMR Biomed* 2014;27(3):312–9.
- [20] Basukala D, Mukundan R, Melzer T, Lim A. Automated segmentation of substantia nigra and red nucleus in quantitative susceptibility mapping images. In: *Proceedings of the 20th international conference on parallel and distributed computing, applications and technologies (PDCAT)*; 2019. p. 375–80.
- [21] Li C, Kao CY, Gore JC, Ding Z. Minimization of region-scalable fitting energy for image segmentation. *IEEE Trans Image Process* 2008;17(10):1940–9.
- [22] Celik T, Tjahjadi T. Unsupervised colour image segmentation using dual-tree complex wavelet transform. *Comput Vis Image Underst* 2010;114(7):813–26.
- [23] Li C, Huang R, Ding Z, Gatenby JC, Metaxas DN, Gore JC. A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI. *IEEE Trans Image Process* 2011;20(7):2007–16.
- [24] Dice LR. Measures of the amount of ecologic association between species. *Ecology* 1945;26(3):297–302.
- [25] McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods* 1996;1(1):30–46.
- [26] An H, Zeng X, Niu T, Li G, Yang J, Zheng L, Zhou W, Liu H, Zhang M, Huang D, Li J. Quantifying iron deposition within the substantia nigra of Parkinson's disease by quantitative susceptibility mapping. *J Neurol Sci* 2018;386:46–52. 2018/03/15/.
- [27] Barbosa JH, Santos AC, Tumas V, Liu M, Zheng W, Haacke EM, Salmon CE. Quantifying brain iron deposition in patients with Parkinson's disease using quantitative susceptibility mapping, R2 and R2\*. *Magn Reson Imaging* 2015;33(5):559–65.
- [28] Du G, Lewis MM, Sica C, He L, Connor JR, Kong L, Mailman RB, Huang X. Distinct progression pattern of susceptibility MRI in the substantia nigra of Parkinson's patients. *Mov Disord* 2018;33(9):1423–31.
- [29] Bergsland N, Zivadinov R, Schweser F, Hagemeier J, Lichter D, Guttuso T. Ventral posterior substantia nigra iron increases over 3 years in Parkinson's disease. *Mov Disord* 2019;34(7):1006–13.
- [30] Kim HG, Park S, Rhee HY, Lee KM, Ryu CW, Rhee SJ, Lee SY, Wang Y, Jahng GH. Quantitative susceptibility mapping to evaluate the early stage of Alzheimer's disease. *Neuroimage Clin* 2017;16:429–38.

**Dibash Basukala** is a Ph.D. student in the Computer Science and Software Engineering Department at the University of Canterbury, Christchurch, New Zealand. His research interests include MRI image segmentation, image feature extraction, and medical image analysis.

**Ramakrishnan Mukundan** is a Professor in the Department of Computer Science and Software Engineering at the University of Canterbury in New Zealand. His primary research interests are in the areas of computer vision, biomedical image analysis, machine learning, and computer graphics. He has authored 3 books, 2 book chapters and over 120 technical papers in various journals and conference proceedings.

**Anthony Lim** received MBChB from University of Otago and FRANZCR from the Royal Australian and New Zealand College of Radiologists. Currently, he works as a consultant neuroradiologist at Canterbury District health Board. His work is related to MRI and Neuroradiology.

**Michael Hurrell** trained in medicine in Auckland, completed a diploma in biomedical engineering at the University of London, then trained in radiology at Christchurch Hospital, becoming a consultant in 1985. He practices as a neuroradiologist and general radiologist, and has examined for the RANZCR, examining in anatomy, pathology, and radiodiagnosis. His research interests include MARS CT, neuroscience and education.

**Ross Keenan** graduated from the University of Otago, New Zealand and completed his Fellowship training in Paediatric Radiology and Neuroradiology. He is a radiologist, and the consultant neuroradiologist at the New Zealand Brain Research Institute, Christchurch, New Zealand.

**John Dalrymple-Alford** is a Professor of Psychology at the University of Canterbury and New Zealand Brain Research Institute, both in Christchurch, New Zealand. His primary interests include neurodegenerative disorders, particularly Parkinson's disease and cognitive correlates, and behavioural neuroscience, particularly diencephalic amnesia and recovery of memory after brain injury.

**Tim Anderson** is a neurologist and Cas Van Der Veer Chair in Parkinson's and Movement Disorders at the University of Otago, Christchurch and Clinical Director of the New Zealand Brain Research Institute. His research interests are biomarkers of cognitive impairment in Parkinson's disease, and eye movements in health and neurodegenerative diseases.

**Daniel Myall** is a Research Fellow at the New Zealand Brain Research Institute. He received a B.Sc. (Hons, 1st Class) in Mathematics from the University of Canterbury and a Ph.D. in Medicine from the University of Otago. He is a Bayesian statistician with a particular focus on Parkinson's and epidemiology.

**Tracy R Melzer** is a Senior Research Fellow in the Department of Medicine, University of Otago, Christchurch, New Zealand, and is the MRI research manager at the New Zealand Brain Research Institute. He is interested in the application of multimodal neuroimaging data, ranging from very preterm birth to diseases of the elderly.