



OPEN ACCESS

EDITED BY

Santiago Perez-Lloret,
National Scientific and Technical Research
Council (CONICET), Argentina

REVIEWED BY

Ruchira Pratihari,
University of South Florida, United States
Dafa Shi,
Second Affiliated Hospital of Shantou
University Medical College, China

*CORRESPONDENCE

Ling Ling Chan
✉ ling2chanSGH@gmail.com
Thomas Welton
✉ thomas_welton@nni.com.sg

RECEIVED 29 April 2024

ACCEPTED 01 August 2024

PUBLISHED 20 August 2024

CITATION

Welton T, Hartono S, Lee W, Teh PY, Hou W,
Chen RC, Chen C, Lim EW, Prakash KM,
Tan LCS, Tan EK and Chan LL (2024)
Classification of Parkinson's disease by deep
learning on midbrain MRI.
Front. Aging Neurosci. 16:1425095.
doi: 10.3389/fnagi.2024.1425095

COPYRIGHT

© 2024 Welton, Hartono, Lee, Teh, Hou,
Chen, Chen, Lim, Prakash, Tan, Tan and Chan.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that
the original publication in this journal is cited,
in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Classification of Parkinson's disease by deep learning on midbrain MRI

Thomas Welton^{1,2*}, Septian Hartono^{1,2}, Weiling Lee³,
Peik Yen Teh³, Wenlu Hou³, Robert Chun Chen^{1,2,3},
Celeste Chen³, Ee Wei Lim¹, Kumar M. Prakash¹, Louis C.
S. Tan^{1,2}, Eng King Tan^{1,2} and Ling Ling Chan^{1,2,3*}

¹National Neuroscience Institute (NNI), Singapore, Singapore, ²Duke-NUS Medical School, Singapore, Singapore, ³Singapore General Hospital, Singapore, Singapore

Purpose: Susceptibility map weighted imaging (SMWI), based on quantitative susceptibility mapping (QSM), allows accurate nigrosome-1 (N1) evaluation and has been used to develop Parkinson's disease (PD) deep learning (DL) classification algorithms. Neuromelanin-sensitive (NMS) MRI could improve automated quantitative N1 analysis by revealing neuromelanin content. This study aimed to compare classification performance of four approaches to PD diagnosis: (1) N1 quantitative "QSM-NMS" composite marker, (2) DL model for N1 morphological abnormality using SMWI ("Heuron IPD"), (3) DL model for N1 volume using SMWI ("Heuron NI"), and (4) N1 SMWI neuroradiological evaluation.

Method: PD patients ($n = 82$; aged 65 ± 9 years; 68% male) and healthy-controls ($n = 107$; 66 ± 7 years; 48% male) underwent 3 T midbrain MRI with T2*-SWI multi-echo-GRE (for QSM and SMWI), and NMS-MRI. AUC was used to compare diagnostic performance. We tested for correlation of each imaging measure with clinical parameters (severity, duration and levodopa dosing) by Spearman-Rho or Kendall-Tao-Beta correlation.

Results: Classification performance was excellent for the QSM-NMS composite marker (AUC = 0.94), N1 SMWI abnormality (AUC = 0.92), N1 SMWI volume (AUC = 0.90), and neuroradiologist (AUC = 0.98). Reasons for misclassification were right-left asymmetry, through-plane re-slicing, pulsation artefacts, and thin N1. In the two DL models, all 18/189 (9.5%) cases misclassified by Heuron IPD were controls with normal N1 volumes. We found significant correlation of the SN QSM-NMS composite measure with levodopa dosing ($\rho = -0.303$, $p = 0.006$).

Conclusion: Our data demonstrate excellent performance of a quantitative QSM-NMS marker and automated DL PD classification algorithms based on midbrain MRI, while suggesting potential further improvements. Clinical utility is supported but requires validation in earlier stage PD cohorts.

KEYWORDS

machine learning, substantia nigra, Nigrosome-1, Parkinson's disease, susceptibility, MRI, iron, neuromelanin

1 Introduction

Parkinson's disease (PD) diagnosis is a major clinical challenge owing to its wide clinical and aetiological heterogeneity, its overlap with other entities, and the lack of reliable *in-vivo* biomarkers. The primary neuropathological hallmark of PD is the progressive loss of dopaminergic (DA) neurons in the iron-rich substantia nigra pars compacta (SNpc) (Dickson, 2012). Nigrosomes 1–5 are located within the SNpc, of which N1 is the largest, and the site of the most sensitive marker of PD pathology histologically (Dickson, 2012). When N1 degeneration occurs, neuromelanin is released and iron is deposited into the extra-cellular space.

Differences in iron within the SNpc can be detected on iron-sensitive MR sequences. On T2*-weighted or susceptibility weighted MRI (SWI), the normal SNpc appears hypointense and the normal N1 appears hyperintense, resulting in the “swallow-tail” sign (Schwarz et al., 2014), which has excellent classification performance for PD patients versus healthy controls (HC; sensitivity=94.6%, specificity=94.4%) (Mahlknecht et al., 2017). Neuromelanin differences can be detected on specialized neuromelanin-sensitive (NMS) fast spin-echo sequences (Blazejewska et al., 2013). However, both approaches require expert visual radiological assessment, and carry the risk of observer-dependent rater bias.

Susceptibility map weighted imaging (SMWI) avoids artifacts induced by phase, and has increased susceptibility contrast and SNR (Gho et al., 2014), allowing a more accurate N1 assessment. Radiological PD classification using SMWI has excellent performance (accuracy=91.8–97.7%) (Liu et al., 2020; Sung et al., 2022). Deep learning (DL) approaches using SMWI can also identify N1 abnormality (Heuron IPD; Heuron Co., Ltd., Seoul, Republic of Korea) and diagnose PD with excellent classification performance (AUC=0.95, Shin et al., 2021). However, this proprietary method has not been independently externally validated and the benefit of a DL approach that automates N1 volume quantification (Heuron NI; Heuron Co., Ltd., Seoul, Republic of Korea) (Jeong et al., 2022) is unknown. The original report on the validation of Heuron IPD was limited by its stringent selection of participants (e.g., according to the PET detection of nigrostriatal degeneration, or the appearance of the N1 on MRI) which, while helpful for model training, may have also inflated the validation AUC (Shin et al., 2021). Secondly, the control sample in the original report comprised individuals with drug-induced Parkinsonism, rather than normal healthy controls (Shin et al., 2021).

Other DL based tools for PD diagnostic classification using MRI have been described in the literature. Several studies have utilised conventional MRI of the cerebrum (e.g., T1-weighted, T2-weighted and FLAIR) to classify PD based on morphological differences, and have achieved accuracies of approximately 90–96% (Basnin et al., 2021; Dhinagar et al., 2021; Camacho et al., 2023, 2024; Mallik et al., 2023). However, atrophy of the cerebrum typically becomes prominent only in the later stages of disease progression, while earlier-stage neuropathology is located in the midbrain (Filippi et al., 2020; Pieperhoff et al., 2022). Fewer studies have investigated the use of midbrain neuropathology from MRI as input to DL algorithms (Huseyn, 2020). Secondly, since morphological changes are nonspecific, studies have utilised advanced MRI techniques such as QSM and NMS MRI of midbrain nuclei in DL tools for PD diagnosis, which are better able to detect PD-specific neuropathological

processes such as iron and neuromelanin content (Shinde et al., 2019; Gaurav et al., 2022a; Wang et al., 2023; Chen et al., 2024).

In this study we evaluated the Heuron IPD and Heuron NI DL models on our database of PD patients with midbrain SMWI. We compared the N1 DL models against an iron-neuromelanin composite model to determine the value-add of NMS in PD diagnosis. We hypothesised that additional NMS provides integral data that would improve the classification performance compared to either N1 DL model alone. We hypothesised that the DL models would demonstrate classification performance of AUC > 0.9, comparable to an experienced neuroradiologist, and comparable to that previously reported by the model developers (Shin et al., 2021).

2 Methods

2.1 Patient population

We used MRI and clinical data from PD patients (OFF-medication state) and age-matched HC, who were recruited from clinics at our tertiary referral centre between 2019 and 2021. PD patients were diagnosed by four neurologists specializing in movement disorders (mean 17.8 years of experience) using the Movement Disorder Society Clinical Diagnostic Criteria for Parkinson's disease (Postuma et al., 2015). Age-matched HC were recruited from the spouses of patients in hospital clinics, health screening and the community, and were absent of neurological conditions. We excluded subjects with MRI contraindications, claustrophobia, known neurological/psychiatric diagnosis other than PD, chronic debilitating medical conditions, or poor cognitive function that would hinder patients' understanding of the study. This study was approved by the local ethics board and written informed consent was obtained from all participants.

2.2 MRI protocols

All MRI data were acquired on the same 3 T MRI system (Siemens Skyra, Erlangen, Germany). We acquired a 3D T2* SWI multi-echo gradient echo sequence with the following parameters: TR 48 ms, TE 13.77/26.39/39 ms, FA 20°, voxel size 0.5 × 0.5 × 1 mm³, 32 slices, duration: 4.15 min. An echo train length of 3 was determined to be an acceptable trade-off between SNR and clinically-feasible acquisition time.

We also acquired an NMS T1-weighted turbo spin echo sequence with the following parameters: TR 938 ms, TE 15 ms, voxel size 0.5 × 0.5 × 3 mm³, 13 slices, duration: 10.42 min. Both sequences were acquired in an oblique-coronal scan plane positioned perpendicular to the midbrain, to improve the N1 in-plane visualization.

2.3 QSM post-processing

Quantitative susceptibility in parts per billion (ppb) was computed from QSM using the STI Suite (Li et al., 2014). Brain extraction based on the magnitude images was performed using the FSL Brain Extraction Tool (BET2). Phase unwrapping and background field phase removal were performed using the HARPERELLA technique. Regularized k-space inverse filtering was performed on the processed

phase images to generate the initial QSM images. An iterative k-space algorithm was used on the initial QSM images to yield the final mean susceptibility (iron deposition) map (Haacke et al., 2010; Barbosa et al., 2015).

2.4 SMWI post-processing

The SMWI images were reconstructed using the SMWI software (Seoul National University, Seoul, Republic of Korea) (Nam et al., 2017) from the multi-echo GRE images as follows: (1) the channel-combined magnitude images were created using the root sum of squares of the multi-channel magnitude images, (2) the channel-combined phase images were created as the mean after correcting for phase offsets of individual channels, (3) the magnitude images from each echo were combined by root sum of squares, (4) the phase images from each echo were unwrapped by Laplacian unwrapping and a frequency calculated per voxel, (5) the background field was removed from the frequency images, (6) the QSM images were reconstructed using the sparse linear equation and least-squares method, (7) a QSM mask was created based on a paramagnetic threshold value, (8) the SMWI was generated as the product of the combined magnitude image and the QSM mask.

2.5 Clinical severity measurements

All participants underwent a clinical motor assessment using the Movement Disorders Society Unified Parkinson's Disease Rating Scale motor part (MDS-UPDRS-III) (Goetz et al., 2008), and the Hoehn and Yahr stage (H&Y) (Hoehn and Yahr, 1967). We also recorded the levodopa equivalent daily dose (LEDD) and disease duration (age at MRI minus age at diagnosis) for the PD group.

2.6 QSM-NMS composite heuristic measure

We formed a heuristic measure for PD classification by combining information from QSM and NMS scans with the following steps. (1) Blinded manual segmentation of the whole SN region using MRICroGL (University of South Carolina, Columbia, SC) on QSM and NMS separately, on three consecutive slices by a neuroradiologist. The slices were selected by inspecting the images in cranio-caudal direction and identifying the first slice whereby the red nuclei were barely or no-longer visible, and the two inferior consecutive slices. (2) Thresholding of these SN volumes on QSM and NMS images separately, for low susceptibility and high neuromelanin content as previously described (Schwarz et al., 2011; Kim et al., 2018; Hartono et al., 2023). The threshold selected was that which maximized the difference between PD and control groups. (3) In the NMS images, calculation of the ratio of the 90th to the 10th percentile of the NMS signal within the SN mask defined the "NMS contrast range." (4) The "QSM-NMS composite" score was defined as the product of the QSM and NMS-based volumes, and the NMS contrast range. We determined this formula based on the preliminary observations that PD patients had smaller SN volumes in QSM and NMS images, and more narrow ranges of contrast in NMS images. Thus, lower values indicated a

higher likelihood of PD and *vice-versa*. We formed an aggregate measure by averaging the left and right-sided values, to increase the SNR and to reduce the multiple testing burden among the set of primary tests.

2.7 Deep learning models

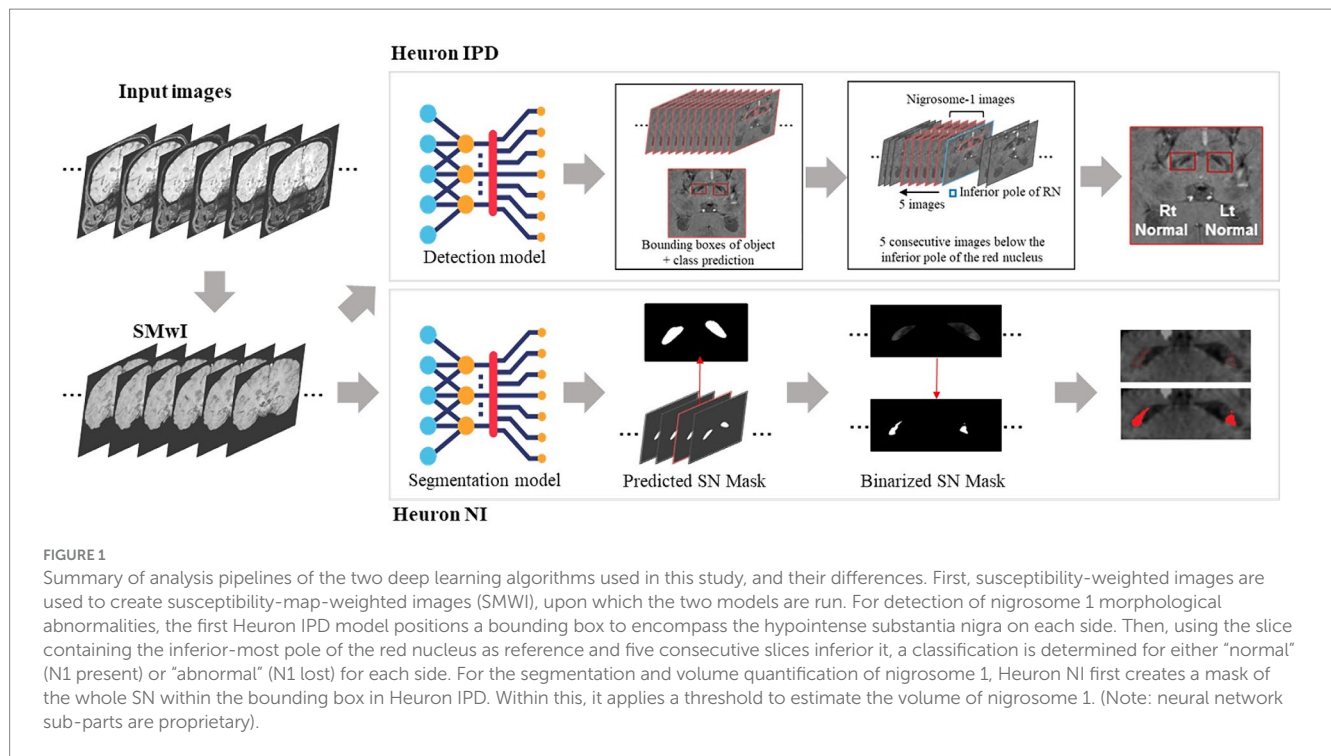
Two proprietary commercial DL models were provided to us (Figure 1; Heuron Co., Ltd., Seoul, Republic of Korea). In "Heuron IPD," five slices were first automatically identified on the SMWI containing the N1 before detection of any abnormality (Shin et al., 2021). Abnormalities were detected using a convolutional deep neural network (CNN), YOLOv3 (Redmon and Farhadi, 2018), to detect morphological abnormality of the N1 region from the SMWI images. Heuron IPD returned a binary classification of "Normal" or "Abnormal" (Shin et al., 2021). "Heuron NI" automatically detected and segmented hyperintensities in the same N1-containing cuts on SMWI (Jeong et al., 2022), and returned the volume of the N1 in mm³. Heuron NI utilizes SparseInst for segmentation of the SN region (Cheng et al., 2022), and is based on a fully-convolutional encoder-decoder architecture, which includes backbone, context-encoder, and decoders to create instance activation maps. The model was trained using ResNet as the backbone, AdamW as the optimizer (with learning rate 5e-5) and a batch size of 16. The training of the model involved focal, dice, and binary cross entropy loss functions. Data augmentation was used to re-scale, and adjust brightness of the input data. Both programs provided left and right hemispheric results, which we analysed separately. For Heuron IPD, we also aggregated the left and right sided data by classifying subjects only as "Normal" if both the left and right N1 were "Normal," and otherwise as "Abnormal." For Heuron NI, we aggregated the left and right sided data by averaging the left and right volumes.

2.8 Neuroradiologist assessment

An experienced neuroradiologist (25 years) performed the assessment of SMWI while blinded to the subject status. Each side was rated as normal (clear visualization of N1), or abnormal (complete or suspected N1 loss). A subject was classified as normal when both sides were rated normal (Sung et al., 2022). Real time SMWI image reformatting was performed as needed (e.g., symmetry alignment, axial bicommissural planar or its orthogonal review) to improve the clarity for assessment as per routine clinical workflow.

2.9 Statistical analysis

Statistical analysis was performed using SPSS version 26 (IBM SPSS Statistics, IBM Corp, Armonk, NY). Continuous imaging measures were right-skewed in both groups and so we report the median and inter-quartile range statistics and used non-parametric tests. The clinical and demographic measures were approximately normally distributed so we report the mean and SD, and apply parametric tests. We performed ROC curve analysis to calculate the AUC. For the models with continuous outcomes (Heuron NI volume, QSM-NMS composite, QSM- and NMS-based volumes, NMS contrast



range), we binarized the data using the Youden Index. We compared the AUCs of Heuron IPD, Heuron NI, neuroradiologist, and the QSM-NMS composite. Correlation of the continuous imaging measures in (1) each hemisphere separately and (2) averaged between hemispheres, with clinical parameters in both groups combined, and in the PD group only was performed using the Spearman rank correlation and Kendall Tau Beta. Multiple comparisons were controlled using the Bonferroni method ($\alpha = 0.05/40 = 0.00125$).

3 Results

3.1 Sample characteristics

Our final sample comprised data from 189 participants including 82 PD (aged 65 ± 9 years; 68% male) and 107 HC (aged 66 ± 7 years; 48% male; Table 1). Patients had mild disease (MDS-UPDRS-III = 31 (20–38); H&Y stage = 2 (1–3); LEDD = 375 mg (250–564), disease duration = 4.8 (1.44–8.70) years). The PD and HC groups differed significantly on all quantitative imaging measures, both as whole groups (Table 1; $p < 0.001$) and when split by hemisphere (Table 2; $p < 0.001$). Heuron NI was unable to process one PD case due to a severe pulsation artifact, thus the sample size for Heuron NI is 188 (Supplementary Figures S1, S2, showing MRI images of the pulsation artefact and a flowchart of subject inclusion).

3.2 Classification performance

The full classification performance results are summarized in Table 3. The QSM-NMS iron-neuromelanin composite measure showed an excellent classification performance (AUC = 0.94, accuracy = 89%, sensitivity = 94%, specificity = 86%; Figure 2),

comparable to an experienced neuroradiologist (AUC = 0.98). The false positives (15/189; 7.9%) had smaller NMS- ($p < 0.001$) and QSM-based volumes ($p = 0.018$) than other HCs (Mann-Whitney U tests). Accordingly, the false negatives (5/189; 2.6%) had larger Heuron NI ($p = 0.021$), NMS- ($p = 0.008$) and QSM-based volumes ($p < 0.001$), and iron-neuromelanin composite scores ($p < 0.001$; Mann-Whitney U tests).

Heuron IPD could also classify PD patients with excellent performance (AUC = 0.92, accuracy = 90% sensitivity = 100%, specificity = 83%). The 18/189 (9.5%) cases that were incorrectly classified (all false positives) had smaller left Heuron NI volumes ($p = 0.010$) and lower iron-neuromelanin composite scores ($p = 0.039$; Mann-Whitney U tests) than correctly-classified HCs.

Heuron NI (N1 volume) classified PD patients with moderate performance (AUC = 0.90, accuracy = 85%, sensitivity = 84%, specificity = 85%). There were both false positive (16/188; 8.5%) and false negative (13/188; 7.0%) classifications which, together, did not differ from the correctly classified cases in any demographic, clinical or imaging measure (Mann-Whitney U test; all $p > 0.05$). The false positives alone had smaller QSM-based volumes ($p < 0.001$) and lower iron-neuromelanin composite scores ($p < 0.001$) than the correctly-classified HCs (Mann-Whitney U tests). Conversely, the false negatives had larger QSM- ($p < 0.001$), and NMS-based volumes ($p = 0.008$), and iron-neuromelanin composite scores ($p < 0.001$) compared to the correctly classified PD patients (Mann-Whitney U tests). Thus, the falsely-classified subjects were outliers with regard to volume in their respective groups but could not be distinguished on demographic or clinical measures.

As a benchmark, our experienced neuroradiologist performed a classification using the SMwI alone with excellent performance (AUC = 0.98, accuracy = 97%, sensitivity = 99%, specificity = 96%), with the most notable difference from the DL and quantitative approaches being the high specificity. A neuroradiologist visual *post-hoc*

TABLE 1 Demographic and clinical information, and quantitative substantia nigra measurements for Parkinson’s disease and control groups, compared between groups.

	Parkinson’s disease, <i>n</i> = 82	Healthy controls, <i>n</i> = 107	Group difference, <i>p</i> -value
Age, years ^a	65.00 (9.29)	65.51 (6.57)	0.656
Gender ^b	Male = 56 (68%) Female = 26 (32%)	Male = 51 (48%) Female = 56 (52%)	0.005 **
MDS-UPDRS-III ^c	31 (20–38)	2 (0–6)	<0.001 **
H&Y stage ^b	0 = 0 (0%) 1 = 5 (6%) 2 = 68 (83%) 3 = 9 (11%)	0 = 107 (100%) 1 = 0 (0%) 2 = 0 (0%) 3 = 0 (0%)	<0.001 **
LEDD, mg ^c	375 (250–564)	–	–
Disease duration, years between diagnosis and MRI ^c	4.83 (1.44–8.70)	–	–
Heuron NI volume, mm ³ ^c	3.00 (1.35–5.44)	14.00 (8.94–20.00)	<0.001 **
QSM-based volume, mm ³ ^c	27.50 (15.50–45.00)	98.00 (56.00–142.00)	<0.001 **
NMS-based volume, mm ³ ^c	16.00 (6.00–27.00)	51.00 (25.00–108.00)	<0.001 **
NMS contrast range ^c	1.139 (1.126–1.156)	1.173 (1.163–1.188)	<0.001 **

MDS-UPDRS-III, Movement Disorders Society Unified Parkinson Disease Rating Scale Part III; H&Y, Hoehn & Yahr; LEDD, levodopa equivalent daily dose; QSM, quantitative susceptibility mapping; NMS, neuromelanin-sensitive.

***p* < 0.01.

^aMean (standard deviation); groups compared by independent samples t-test.

^bFrequency (percent); groups compared by Chi-square test.

^cMedian (inter-quartile range); groups compared by Mann–Whitney U test.

TABLE 2 Descriptive statistics for quantitative substantia nigra MRI measurements, split by hemisphere and compared between Parkinson’s disease and healthy control groups.

	Parkinson’s disease, <i>n</i> = 82	Healthy controls, <i>n</i> = 107	Group difference, <i>p</i> -value
Heuron NI volume, mm ³ ^a			
Left	2.75 (1.25–4.87)	14.62 (9.12–20.38)	<0.001 **
Right	2.38 (0.88–6.19)	11.88 (7.50–19.50)	<0.001 **
QSM-based volume, mm ³ ^a			
Left	25.00 (12.00–47.00)	95.00 (58.00–114.00)	<0.001 **
Right	27.50 (14.75–47.00)	93.00 (52.00–139.00)	<0.001 **
NMS-based volume, mm ³ ^a			
Left	23.00 (7.50–47.75)	106.00 (54.00–168.00)	<0.001 **
Right	16.50 (5.75–29.25)	51.00 (25.00–108.00)	<0.001 **
NMS contrast range ^a			
Left	1.140 (1.126–1.156)	1.174 (1.163–1.192)	<0.001 **
Right	1.158 (1.136–1.172)	1.193 (1.175–1.211)	<0.001 **

QSM, quantitative susceptibility mapping; NMS, neuromelanin-sensitive.

***p* < 0.01.

^aMedian (inter-quartile range); groups compared by Mann–Whitney U test.

investigation of subjects misclassified by DL models identified (1) motion, cardiac pulsation (Supplementary Figure S1, showing pulsation artefact) and streaking artefacts, (2) bilaterally thin N1, (3) through-plane re-slicing which reduces signal and compounds aforesaid factors, (4) right–left alignment asymmetry from head tilt, (5) inappropriate slice selection for N1 segmentation, and (6) frequent overestimation of the N1 mask, as likely factors contributing to the false classifications of the DL models (Figure 3).

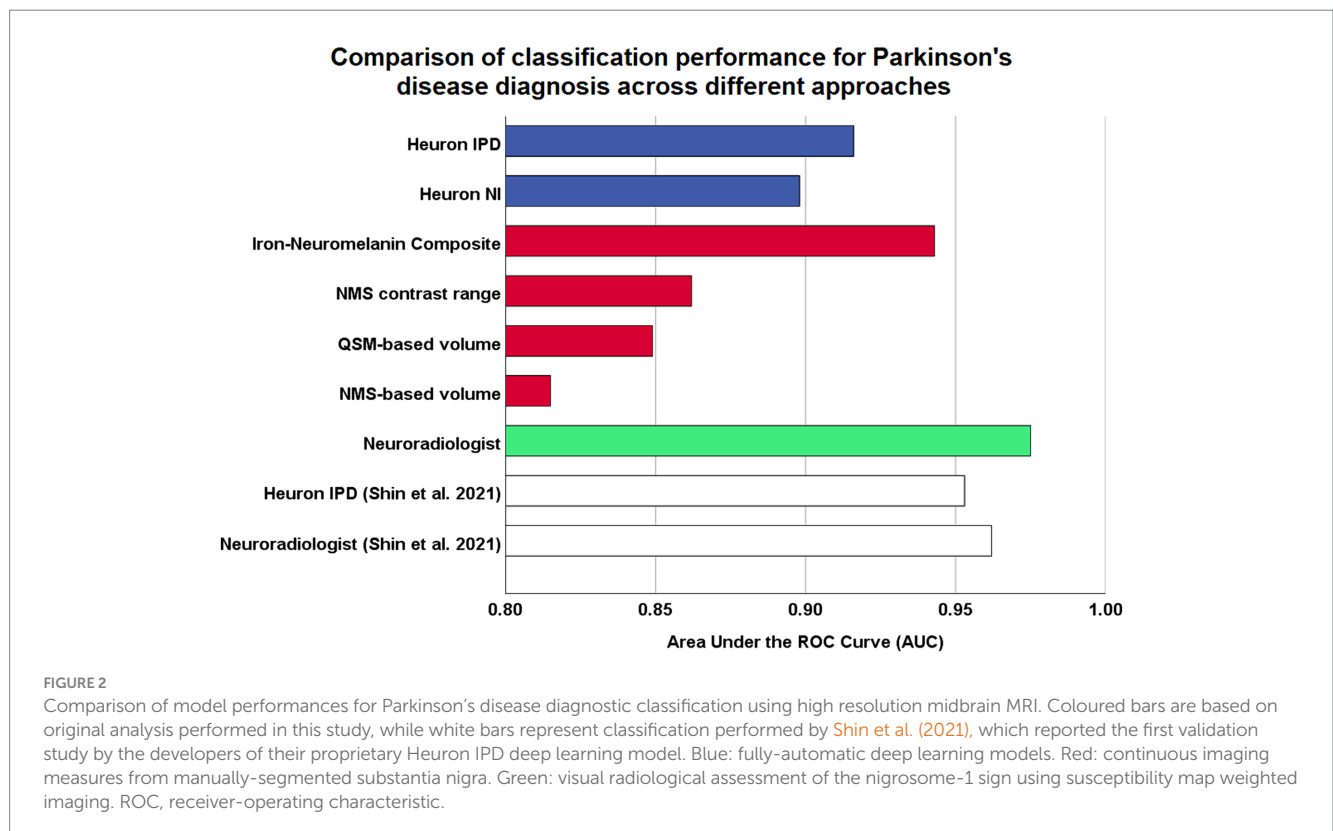
3.3 Hemispheric differences

The left hemisphere had consistently better performance than the right across all models except for Heuron IPD. The left–right discrepancy in AUCs for Heuron NI volume, NMS contrast range, NMS-based volume, QSM-based volume, iron-neuromelanin composite and neuroradiologist was 0.059, 0.005, 0.026, 0.052, 0.030 and 0.002, respectively. Our PD cohort was 100% right-handed, so

TABLE 3 Results of the binary Parkinson’s disease versus healthy control classification by each model based on bilateral and single hemispheric findings.

Model	AUC	Accuracy (%)	Sensitivity (%)	Specificity (%)	Youden Index	Optimal cutoff
Heuron IPD	0.916	90.48	100.00	83.18	–	–
Left	0.912	90.48	97.56	85.05	–	–
Right	0.930	92.06	100.00	85.98	–	–
Heuron NI volume	0.898	84.57	83.95	85.05	0.69	6.88
Left	0.905	84.57	83.95	85.05	0.69	7.19
Right	0.846	78.72	81.48	76.64	0.58	7.19
QSM-NMS Composite	0.943	89.42	93.90	85.98	0.80	2178.62
Left	0.898	84.66	85.37	84.11	0.69	3297.83
Right	0.868	79.89	89.02	72.90	0.61	2110.04
NMS contrast range	0.862	80.95	84.15	78.50	0.63	1.1590
Left	0.837	80.42	79.27	81.31	0.60	1.1570
Right	0.832	78.31	76.83	79.44	0.55	1.1723
NMS-based volume	0.815	76.72	71.95	80.37	0.51	23.50
Left	0.832	78.31	79.27	77.57	0.56	53.50
Right	0.806	74.60	67.07	80.37	0.46	23.50
QSM-based volume	0.849	78.84	81.71	76.64	0.59	52.75
Left	0.856	80.42	76.83	83.18	0.60	49.00
Right	0.804	77.25	80.49	74.77	0.56	57.50
Neuroradiologist	0.975	97.35	98.78	96.26	–	–
Left	0.960	96.30	95.12	97.20	–	–
Right	0.958	96.30	93.90	98.13	–	–

AUC, area under the receiver operating characteristic curve; NMS, neuromelanin-sensitive; QSM, quantitative-susceptibility mapping.



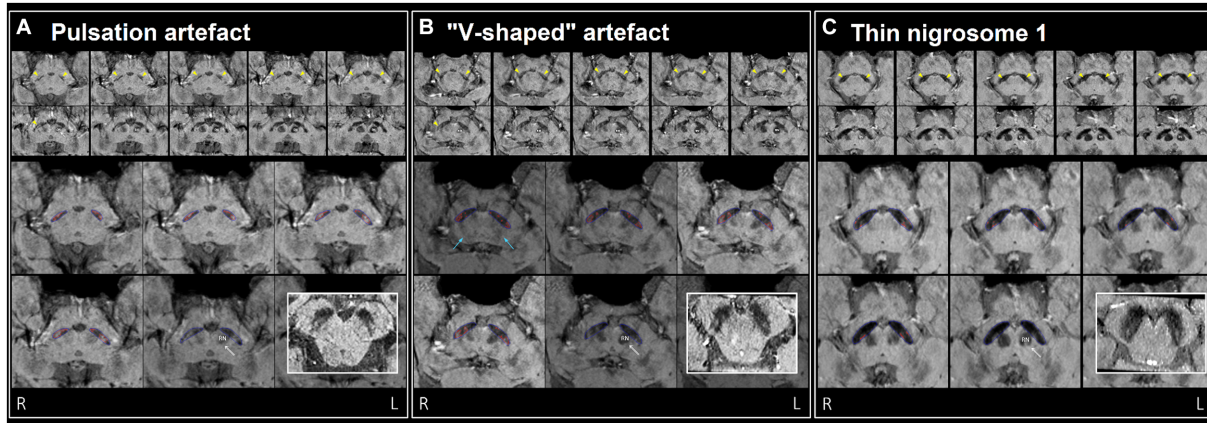


FIGURE 3

Examples of misclassification by DL model(s) in three healthy controls (false positives), **(A–C)** visually assessed as normal by neuroradiologist. Top two rows: ten consecutive 0.5 mm re-sliced caudo-cranial susceptibility map weighted (SMWI) output images from Heuron IPD, orientated perpendicular to the midbrain, showing the hyperintense nigrosome-1, N1 (yellow arrows) within the hypointense substantia nigra on cuts inferior to the red nucleus (marked "RN" on the left). Bottom two rows: magnified Heuron NI output images demonstrating hyperintensities outlined in red within the hypointense substantia nigra (blue) on four consecutive caudo-cranial cuts inferior to that containing the left red nucleus (RN) indicated by a double arrow. Bold inset image: intact N1 ("swallow-tail" sign) visualized as a dorsolateral hyperintensity in all three healthy controls on axial images reformatted parallel to the bi-commissural plane, providing a confirmatory alternative imaging perspective. **(A)** Pulsation artefacts from in-plane ambient cisternal arterial loops in this 47-year-old healthy male and slight right-left alignment asymmetry (tilted head - unequal red nuclei) could contribute to the "Abnormal" label of the right N1 on Heuron IPD. However, volume outputs on Heuron NI were normal. **(B)** 67-year-old healthy male with dark V-shaped (blue arrows) artefacts superimposed across the substantia nigra could impair N1 detection, and result in bilaterally "Abnormal" labels on Heuron IPD. Again, bilateral volume outputs on Heuron NI were normal. **(C)** Bilaterally skinny but distinct N1 in this 58-year-old female healthy control were labeled "Abnormal" by Heuron IPD. Volume outputs on Heuron NI were abnormally low; ideally, segmentation could have been automated on more inferior cuts after clearing both red nuclei.

TABLE 4 Correlation (correlation coefficient, *p*-value) of clinical severity, levodopa dosing and disease duration with quantitative substantia nigra measurements separately in each hemisphere in the Parkinson's disease group only.

	MDS-UPDRS-III ^a	H&Y stage ^b	LEDD ^a	Disease duration ^a
Heuron NI volume				
Left	-0.117, 0.308	0.001, 0.995	-0.040, 0.722	0.005, 0.961
Right	-0.037, 0.745	0.126, 0.166	-0.016, 0.887	0.028, 0.803
QSM-based volume				
Left	-0.089, 0.437	-0.076, 0.397	-0.208, 0.062	-0.127, 0.260
Right	0.026, 0.819	0.141, 0.119	0.018, 0.874	0.106, 0.346
NMS-based volume				
Left	-0.149, 0.189	0.015, 0.868	-0.333, 0.002 *	-0.207, 0.064
Right	-0.061, 0.595	-0.035, 0.703	-0.202, 0.071	-0.020, 0.856
NMS contrast range				
Left	-0.243, 0.031 *	-0.027, 0.764	-0.268, 0.015 *	-0.215, 0.054
Right	-0.171, 0.133	-0.051, 0.572	-0.229, 0.040 *	-0.085, 0.449
Iron-neuromelanin composite				
Left	-0.230, 0.041 *	-0.030, 0.740	-0.387, 0.0004 **	-0.211, 0.059
Right	-0.134, 0.240	0.043, 0.632	-0.231, 0.038 *	-0.005, 0.964

MDS-UPDRS-III, Movement Disorders Society Unified Parkinson Disease Rating Scale Part III; H&Y, Hoehn & Yahr; LEDD, levodopa equivalent daily dose; QSM, quantitative-susceptibility mapping; NMS, neuromelanin-sensitive. **significant after multiple comparison correction, $\alpha < 0.00125$.

* $p < 0.05$.

^aSpearman correlation, reported as ρ .

^bKendall Tau Beta correlation, reported as τ_b .

we could not further evaluate the impact of handedness. The side of initial symptom onset for PD patients was left for $n = 21$, right for $n = 49$, symmetrical for $n = 7$ and unknown for $n = 5$. However, comparison of the left- and right-onset patients on the continuous imaging measures showed no significant effects (Mann-Whitney U test).

3.4 Correlation of imaging measures with clinical severity

Correlation of clinical motor symptom severity (MDS-UPDRS-III, H&Y stage) with the continuous imaging measures when both the PD and HC groups were included showed more severe disease correlating with smaller Heuron NI-, QSM- and NMS-based volumes and lower NMS contrast (Spearman $\rho < -0.476$, $p < 6.75e^{-12}$).

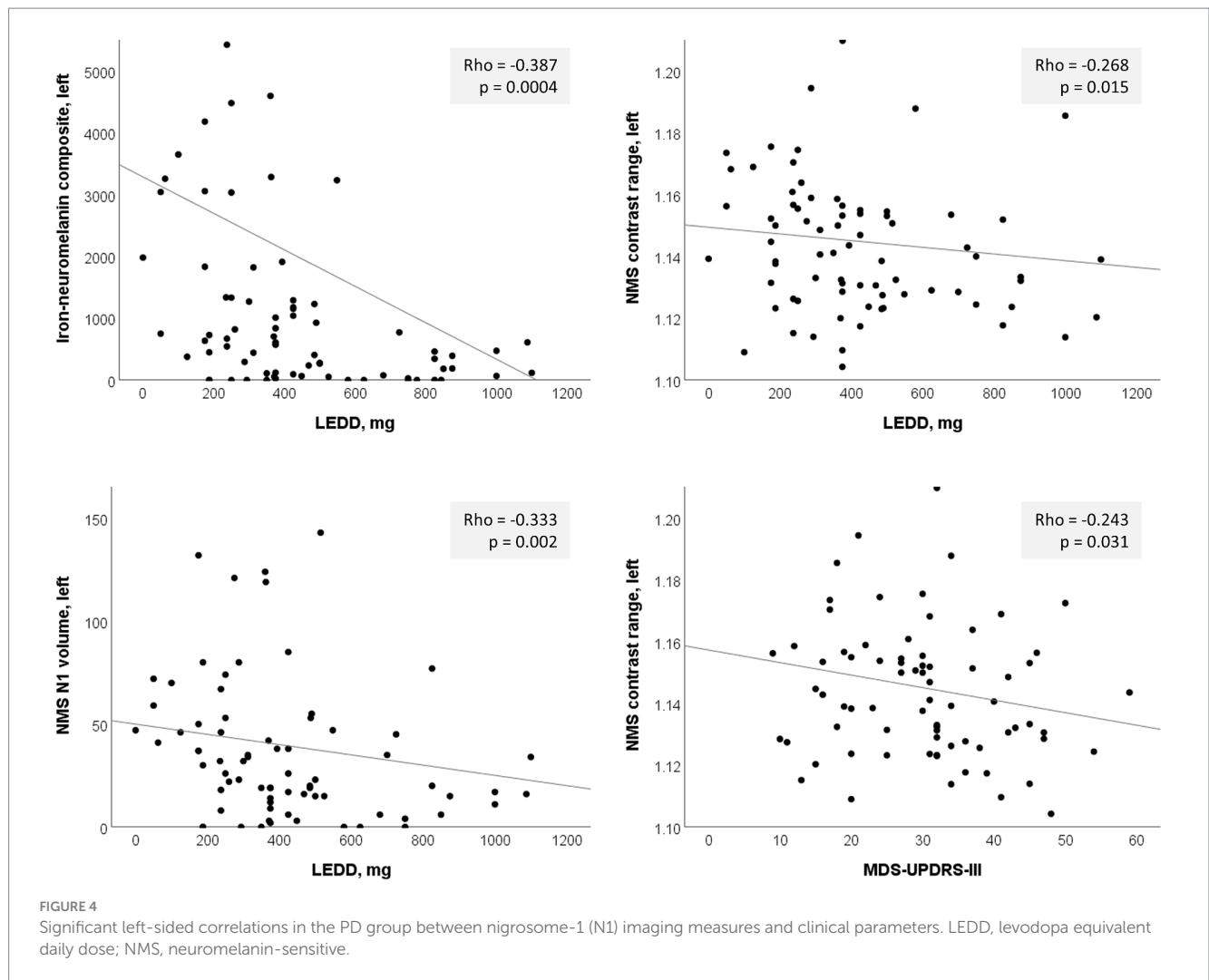
Based on the hemispheric differences we observed, we explored *post-hoc* correlations using individual hemispheric imaging measures in the PD group only, correcting for multiple comparisons using the Bonferroni method ($\alpha = 0.05/40 = 0.00125$). Since we included the PD group only, we also tested for correlation with dosing (LEDD) and disease duration. This confirmed the predominant usefulness of left-sided imaging measures, with mostly left-sided results having $p < 0.05$ (Table 4; Figure 4). Correlation of the continuous imaging measures with severity (MDS-UPDRS-III, H&Y stage) and disease duration was weaker than that with levodopa dosing (LEDD), and was stronger for NMS-based than QSM-based measures. Correlation of the iron-neuromelanin composite and LEDD in the left hemisphere remained significant after multiple comparison correction (Spearman $\rho = -0.303$, $p = 0.006$).

Supplementary Table S1 shows the results of the correlation analyses between the left and right averaged continuous imaging measures and clinical severity for the PD group, which concur with the single-hemisphere results.

4 Discussion

The use of automated tools to supplement PD diagnosis is an ongoing important area of research. Recent progress has focused on classification by characterizing N1 using MRI contrasts sensitive to magnetic susceptibility (iron) and neuromelanin (Shin et al., 2021; Sung et al., 2021; Jokar et al., 2023). We compared classification performance of an iron-neuromelanin composite measure, Heuron IPD and Heuron NI (DL models based on SMWI) against that of an experienced neuroradiologist, to determine the potential value-add of NMS MRI and the independent external validity of the DL models. We demonstrated good value of a combined iron-neuromelanin (QSM-NMS) marker. We found excellent performance for each model, which was comparable to the radiologist. These results mark the first independent external validation of a method for automatic PD classification based on SMWI, supporting its efficacy, while suggesting further improvements that could be made.

An iron-neuromelanin marker had excellent classification performance (AUC = 0.94) exceeding that of either DL model alone, and similar to a recent study using an automated SN template approach (AUC = 0.95) (Jokar et al., 2023). Our method had better performance than other approaches to SN NMS classification; for example, automated NMS quantification (AUC = 0.83) (Gaurav et al., 2022b), and was similar to others using manual segmentation on QSM (AUC = 0.96) (Cheng et al., 2019). NMS-based MRI measures,



including the iron-neuromelanin marker, had the strongest correlations with clinical severity and dosing than other quantitative imaging measures. Our marker was based on manual SN segmentation but, nonetheless, suggests a benefit to combining QSM and NMS in future DL approaches for PD diagnosis.

Heuron IPD also achieved excellent classification performance, confirming our hypothesis and supporting its external validity. The AUC of Heuron IPD was 0.92, while its counterparts in similar studies reported AUC = 0.95 (Shin et al., 2021; Jokar et al., 2023). These models were first published in 2021 (Shin et al., 2021) and were trained on a Korean cohort of PD patients and HCs but had not been independently externally validated. Our cohort is also East-Asian and both have similar disease severity (H&Y stage = 2). However, the classification performance could have been limited by our cohort's younger age (65 versus 71 years). The training cohort was selected based on dopamine transporter DAT scan status and MRI-appearance of the N1 whereas ours was not, so our analysis may represent a more ecologically-valid scenario since DAT scan is not always available. A further important test will be to apply these methods to undiagnosed suspected PD and samples which have not been filtered based on neurodegeneration, N1 structure, or presence of

other neurological or psychiatric conditions. Finally, while each model had comparable AUC to the neuroradiologist, the number of false positives was notably greater. This was common across models but, in general, classification should ideally err on the side of false positives rather than false negatives.

We found that the performance of Heuron NI (AUC = 0.90) was less than that of Heuron IPD. Misclassified cases generally showed motion or pulsation artefacts, intact but thin N1, right-left alignment asymmetry, or reduced signal secondary to re-slicing of the data through-plane, which could confound automated N1 detection. Additional steps to address these could improve classification. The left hemisphere was better for classification than the right hemisphere, and was the only hemisphere to have any significant correlation (after multiple comparison) with levodopa dosing. This may be explained by the tendency for symptoms to first occur on the dominant side (most often right), and thus predominance of left-sided SN neuropathology due to the decussation of cortico-pontine fibres. Concordance of expected side of disease pathology and imaging abnormality serves to validate the imaging approaches. The significant correlation with dosing, but not with severity, suggests structural brain alterations with medication use. Other studies identified relationships between some sub-scores of the MDS-UPDRS-III and

manually-segmented N1/SN on T2-weighted MRI (Fu et al., 2016) but not the MDS-UPDRS-III as a whole.

Clinical demand in radiology for inclusion of high-resolution midbrain imaging in brain MRI orders for evaluation of Parkinsonism is on the rise with increasing availability, evidence of good diagnostic performance (Kuya et al., 2016; Sung et al., 2019), varied clinical presentations (Parkinson's Disease Society of the United Kingdom, 2019) and complex co-morbidities in an aging population (Bloem et al., 2020). For example, an intact N1 and congruous quantitative SN measures in the presence of silent extra-nigral vascular pathology may be useful clinical decision support tools indicating that levodopa should be sparingly prescribed. The principal application for this technology in a clinical workflow is to distinguish patients who have overt N1-sign loss. This could facilitate filtering of cases for reporting between general or junior neuroradiologists and senior neuroradiologists based on case difficulty. Ideally, such tools should be incorporated into an automatic pipeline to not require additional steps, and should present the results directly on a clinical workstation, which requires regulatory approval (Choy et al., 2018). These could also be used as adjunct teaching tools to train radiologists unfamiliar with midbrain N1 assessment.

Future studies should apply SMWI-based DL models in earlier-stage or prodromal PD cohorts, and attempt to classify PD from other Parkinsonisms such as essential tremor (Perez Akly et al., 2019) which may be an early stage PD misdiagnosis (Welton et al., 2021). This approach should also be tested in non-East-Asian cohorts. A limitation is that SMWI requires a specific multi-echo acquisition (Nam et al., 2017) not part of routine clinical neuroimaging protocols. Technologist training is needed for accustomisation to anatomical landmarks for accurate 3D slab placement, as right-left symmetry alignment of the sub-nuclear structures on high resolution SMWI is sensitive to head tilt.

Strengths of our study include the independent, external nature of our validation of DL models. This is important because the cohort used for validation in the original report (Shin et al., 2021) included PD patients based on their PET or MRI status, which could have increased the reported AUC. Our comparisons to midbrain QSM and NMS MRI are original, and this enhanced neuroimaging evaluation of PD by yielding significant correlations with disease severity measures. This is noteworthy for its potential to objectively monitor disease progression compared to QSM-only approaches.

Our data show that automated algorithms, and an iron-NMS marker to augment radiologists' decision-making for PD diagnosis are highly accurate. The DL models can be further improved by incorporation of NMS MRI information, identification of artefacts, combination of data across models, hemispheric information, automatic re-slicing, and further training on other cohorts. There is a potential role for this approach in future clinical workflows, especially to support non-expert radiologists.

Data availability statement

The data analyzed in this study is available from the corresponding author on reasonable request. Requests to access these datasets should be directed to LC, ling2chanSGH@gmail.com.

Ethics statement

The studies involving humans were approved by Singapore Health Services Centralised Institutional Review Board. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

TW: Investigation, Methodology, Formal analysis, Writing – review & editing, Writing – original draft. SH: Data curation, Formal analysis, Investigation, Methodology, Software, Project administration, Writing – review & editing. WL: Investigation, Writing – review & editing. PT: Investigation, Writing – review & editing. WH: Investigation, Writing – review & editing. RC: Investigation, Writing – review & editing. CC: Investigation, Writing – review & editing. EL: Investigation, Writing – review & editing. KP: Investigation, Writing – review & editing. LT: Investigation, Writing – review & editing. ET: Investigation, Writing – review & editing. LC: Conceptualization, Funding acquisition, Data curation, Investigation, Methodology, Software, Resources, Validation, Visualization, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This project was funded by the Singapore National Medical Research Council grant numbers: NMRC/CSASI/20nov-0008 and NMRC/CSA/INV2017.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnagi.2024.1425095/full#supplementary-material>

References

- Barbosa, J. H. O., Santos, A. C., Tumas, V., Liu, M., Zheng, W., Haacke, E. M., et al. (2015). Quantifying brain iron deposition in patients with Parkinson's disease using quantitative susceptibility mapping, R2 and R2*. *Magn. Reson. Imaging* 33, 559–565. doi: 10.1016/j.mri.2015.02.021
- Basnin, N., Nahar, N., Anika, F. A., Hossain, M. S., and Andersson, K. (2021). “Deep learning approach to classify Parkinson's disease from MRI samples” in *BI - Brain informatics*. eds. M. Mahmud, M. S. Kaiser, S. Vassanelli, Q. Dai and N. Zhong (Cham: Springer International Publishing), 536–547.
- Blaziejewska, A. I., Schwarz, S. T., Pitiot, A., Stephenson, M. C., Lowe, J., Bajaj, N., et al. (2013). Visualization of nigrosome 1 and its loss in PD: pathoanatomical correlation and in vivo 7 T MRI. *Neurology* 81, 534–540. doi: 10.1212/WNL.0b013e31829e6fd2
- Bloem, B. R., Henderson, E. J., Dorsey, E. R., Okun, M. S., Okubadejo, N., Chan, P., et al. (2020). Integrated and patient-centred management of Parkinson's disease: a network model for reshaping chronic neurological care. *Lancet Neurol.* 19, 623–634. doi: 10.1016/S1474-4422(20)30064-8
- Camacho, M., Wilms, M., Almgren, H., Amador, K., Camicioli, R., Ismail, Z., et al. (2024). Initiative, exploiting macro- and micro-structural brain changes for improved Parkinson's disease classification from MRI data. *NPJ Park. Dis.* 10:43. doi: 10.1038/s41531-024-00647-9
- Camacho, M., Wilms, M., Mouches, P., Almgren, H., Souza, R., Camicioli, R., et al. (2023). Explainable classification of Parkinson's disease using deep learning trained on a large multi-center database of T1-weighted MRI datasets. *NeuroImage Clin.* 38:103405. doi: 10.1016/j.nicl.2023.103405
- Chen, H., Liu, X., Luo, X., Fu, J., Zhou, K., Wang, N., et al. (2024). An automated hybrid approach via deep learning and radiomics focused on the midbrain and substantia nigra to detect early-stage Parkinson's disease. *Front. Aging Neurosci.* 16:1397896. doi: 10.3389/fnagi.2024.1397896
- Cheng, T., Wang, X., Chen, S., Zhang, W., Zhang, Q., Huang, C., et al. (2022). Sparse instance activation for real-time instance segmentation, in: 2022 IEEE/CVF Conf. Comput. Vis. Pattern Recognit, pp. 4423–4432. doi: 10.1109/CVPR52688.2022.00439
- Cheng, Z., Zhang, J., He, N., Li, Y., Wen, Y., Xu, H., et al. (2019). Radiomic features of the Nigrosome-1 region of the substantia nigra: using quantitative susceptibility mapping to assist the diagnosis of idiopathic Parkinson's disease. *Front. Aging Neurosci.* 11, 1–11. doi: 10.3389/fnagi.2019.00167
- Choy, G., Khalilzadeh, O., Michalski, M., Do, S., Samir, A. E., Panykh, O. S., et al. (2018). Current applications and future impact of machine learning in radiology. *Radiology* 288, 318–328. doi: 10.1148/radiol.2018171820
- Dhinagar, N. J., Thomopoulos, S. I., Owens-Walton, C., Stripelis, D., Ambite, J. L., Ver Steeg, G., et al. (2021). 3D convolutional neural networks for classification of Alzheimer's and Parkinson's disease with T1-weighted brain MRI. *BioRxiv* 5, 1–10. doi: 10.1101/2021.07.26.453903
- Dickson, D. W. (2012). Parkinson's disease and Parkinsonism: neuropathology. *Cold Spring Harb. Pers. Med.* 2, 1–15. doi: 10.1101/cshperspect.a009258
- Filippi, M., Sarasso, E., Piramide, N., Stojkovic, T., Stankovic, I., Basaia, S., et al. (2020). Progressive brain atrophy and clinical evolution in Parkinson's disease. *NeuroImage Clin.* 28:102374. doi: 10.1016/j.nicl.2020.102374
- Fu, K. A., Nathan, R., Dinov, I. D., Li, J., and Toga, A. W. (2016). T2-imaging changes in the Nigrosome-1 relate to clinical measures of Parkinson's disease. *Front. Neurol.* 7, 1–9. doi: 10.3389/fneur.2016.00174
- Gaurav, R., Pyatigorskaya, N., Biondetti, E., Valabrègue, R., Yahia-Cherif, L., Mangone, G., et al. (2022a). Deep learning-based neuromelanin MRI changes of isolated REM sleep behavior disorder. *Mov. Disord.* 37, 1064–1069. doi: 10.1002/mds.28933
- Gaurav, R., Valabrègue, R., Yahia-Cherif, L., Mangone, G., Narayanan, S., Arnulf, I., et al. (2022b). NigraNet: an automatic framework to assess nigral neuromelanin content in early Parkinson's disease using convolutional neural network. *NeuroImage Clin.* 36:103250. doi: 10.1016/j.nicl.2022.103250
- Gho, S.-M., Liu, C., Li, W., Jang, U., Kim, E. Y., Hwang, D., et al. (2014). Susceptibility map-weighted imaging (SMWI) for neuroimaging. *Magn. Reson. Med.* 72, 337–346. doi: 10.1002/mrm.24920
- Goetz, C. G., Tilley, B. C., Shaftman, S. R., Stebbins, G. T., Fahn, S., Martinez-Martin, P., et al. (2008). Movement Disorder Society-sponsored revision of the unified Parkinson's disease rating scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov. Disord.* 23, 2129–2170. doi: 10.1002/mds.22340
- Haacke, E. M., Tang, J., Neelavalli, J., and Cheng, Y. C. N. (2010). Susceptibility mapping as a means to visualize veins and quantify oxygen saturation. *J. Magn. Reson. Imaging* 32, 663–676. doi: 10.1002/jmri.22276
- Hartono, S., Chen, R. C., Welton, T., Sen Tan, A., Lee, W., Teh, P. Y., et al. (2023). Quantitative iron-neuromelanin MRI associates with motor severity in Parkinson's disease and matches radiological disease classification. *Front. Aging Neurosci.* 15:1287917. doi: 10.3389/fnagi.2023.1287917
- Hoehn, M. M., and Yahr, M. D. (1967). Parkinsonism: onset, progression and mortality. *Neurology* 17, 427–442. doi: 10.1212/wnl.17.5.427
- Huseyn, E. (2020). Deep learning based early diagnostics of Parkinson's disease. *ArXiv Prepr.*, 1–14. doi: 10.48550/arXiv.2008.01792
- Jeong, S. Y., Suh, C. H., Heo, H., Shim, W. H., and Kim, S. J. (2022). Current updates and unmet needs of brain MRI-based artificial intelligence software for patients with neurodegenerative diseases in the Republic of Korea. *Investig. Magn. Reson. Imaging* 26, 237–245. doi: 10.13104/imri.2022.26.4.237
- Jokar, M., Jin, Z., Huang, P., Wang, Y., Zhang, Y., Li, Y., et al. (2023). Diagnosing Parkinson's disease by combining neuromelanin and iron imaging features using an automated midbrain template approach. *NeuroImage* 266:119814. doi: 10.1016/j.neuroimage.2022.119814
- Kim, E. Y., Sung, Y. H., Shin, H. G., Noh, Y., Nam, Y., and Lee, J. (2018). Diagnosis of early-stage idiopathic Parkinson's disease using high-resolution quantitative susceptibility mapping combined with histogram analysis in the substantia nigra at 3 T. *J. Clin. Neurol.* 14, 90–97. doi: 10.3988/jcn.2018.14.1.90
- Kuya, K., Shinohara, Y., Miyoshi, F., Fujii, S., Tanabe, Y., and Ogawa, T. (2016). Correlation between neuromelanin-sensitive MR imaging and (123) I-FP-CIT SPECT in patients with parkinsonism. *Neuroradiology* 58, 351–356. doi: 10.1007/s00234-016-1644-7
- Li, W., Avram, A. V., Wu, B., Xiao, X., and Liu, C. (2014). Integrated Laplacian-based phase unwrapping and background phase removal for quantitative susceptibility mapping. *NMR Biomed.* 27, 219–227. doi: 10.1002/nbm.3056
- Liu, X., Wang, N., Chen, C., Wu, P.-Y., Piao, S., Geng, D., et al. (2020). Swallow tail sign on susceptibility map-weighted imaging (SMWI) for disease diagnosing and severity evaluating in Parkinsonism. *Acta Radiol.* 62, 234–242. doi: 10.1177/0284185120920793
- Mahlknecht, P., Krismer, F., Poewe, W., and Seppi, K. (2017). Meta-analysis of dorsolateral nigral hyperintensity on magnetic resonance imaging as a marker for Parkinson's disease. *Mov. Disord.* 32, 619–623. doi: 10.1002/mds.26932
- Mallik, S., Majhi, B., Kashyap, A., Mohanty, S., Dash, S., Li, A., et al. (2023). An improved method for diagnosis of Parkinson's disease using deep learning models enhanced with metaheuristic algorithm. *Res. Sq.* 4, 1–42. doi: 10.21203/rs.3.rs-3387953/v1
- Nam, Y., Gho, S.-M., Kim, D.-H., Kim, E. Y., and Lee, J. (2017). Imaging of nigrosome 1 in substantia nigra at 3T using multiecho susceptibility map-weighted imaging (SMWI). *J. Magn. Reson. Imaging* 46, 528–536. doi: 10.1002/jmri.25553
- Parkinson's Disease Society of the United Kingdom. (2019). Poll finds a quarter of people with Parkinson's are wrongly diagnosed, 2019 UK Park. Audit. Available at: <https://www.parkinsons.org.uk/news/poll-finds-quarter-people-parkinsons-are-wrongly-diagnosed> (Accessed May 4, 2023).
- Perez Akly, M. S., Stefani, C. V., Ciancaglini, L., Bestoso, J. S., Funes, J. A., Bauso, D. J., et al. (2019). Accuracy of nigrosome-1 detection to discriminate patients with Parkinson's disease and essential tremor. *Neuroradiol. J.* 32, 395–400. doi: 10.1177/1971400919853787
- Pieperhoff, P., Südmeyer, M., Dinkelbach, L., Hartmann, C. J., Ferreira, S., Moldovan, A. S., et al. (2022). Regional changes of brain structure during progression of idiopathic Parkinson's disease – a longitudinal study using deformation based morphometry. *Cortex* 151, 188–210. doi: 10.1016/j.cortex.2022.03.009
- Postuma, R. B., Berg, D., Stern, M., Poewe, W., Olanow, C. W., Oertel, W., et al. (2015). MDS clinical diagnostic criteria for Parkinson's disease. *Mov. Disord.* 30, 1591–1601. doi: 10.1002/mds.26424
- Redmon, J., and Farhadi, A. (2018). Yolov3: An incremental improvement. *ArXiv Prepr.*, 1–6. doi: 10.48550/arXiv.1804.02767
- Schwarz, S. T., Afzal, M., Morgan, P. S., Bajaj, N., Gowland, P. A., and Auer, D. P. (2014). The “swallow tail” appearance of the healthy nigrosome – a new accurate test of Parkinson's disease: a case-control and retrospective cross-sectional MRI study at 3T. *PLoS One* 9, –e93814. doi: 10.1371/journal.pone.0093814
- Schwarz, S. T., Rittman, T., Gontu, V., Morgan, P. S., Bajaj, N., and Auer, D. P. (2011). T1-weighted MRI shows stage-dependent substantia nigra signal loss in Parkinson's disease. *Mov. Disord.* 26, 1633–1638. doi: 10.1002/mds.23722
- Shin, D. H., Heo, H., Song, S., Shin, N.-Y., Nam, Y., Yoo, S.-W., et al. (2021). Automated assessment of the substantia nigra on susceptibility map-weighted imaging using deep convolutional neural networks for diagnosis of idiopathic Parkinson's disease. *Parkinsonism Relat. Disord.* 85, 84–90. doi: 10.1016/j.parkreidis.2021.03.004
- Shinde, S., Prasad, S., Saboo, Y., Kaushick, R., Saini, J., Pal, P. K., et al. (2019). Predictive markers for Parkinson's disease using deep neural nets on neuromelanin sensitive MRI. *NeuroImage Clin.* 22:101748. doi: 10.1016/j.nicl.2019.101748
- Sung, Y. H., Kim, J.-S., Yoo, S.-W., Shin, N.-Y., Nam, Y., Ahn, T.-B., et al. (2022). A prospective multi-centre study of susceptibility map-weighted MRI for the diagnosis of neurodegenerative parkinsonism. *Eur. Radiol.* 32, 3597–3608. doi: 10.1007/s00330-021-08454-z

Sung, Y. H., Lee, J., Nam, Y., Shin, H.-G., Noh, Y., Hwang, K. H., et al. (2019). Initial diagnostic workup of Parkinsonism: dopamine transporter positron emission tomography versus susceptibility map-weighted imaging at 3T. *Parkinsonism Relat. Disord.* 62, 171–178. doi: 10.1016/j.parkreldis.2018.12.019

Sung, Y. H., Noh, Y., and Kim, E. Y. (2021). Early-stage Parkinson's disease: abnormal nigrosome 1 and 2 revealed by a voxelwise analysis of neuromelanin-sensitive MRI. *Hum. Brain Mapp.* 42, 2823–2832. doi: 10.1002/hbm.25406

Wang, Y., He, N., Zhang, C., Zhang, Y., Wang, C., Huang, P., et al. (2023). An automatic interpretable deep learning pipeline for accurate Parkinson's disease diagnosis using quantitative susceptibility mapping and T1-weighted images. *Hum. Brain Mapp.* 44, 4426–4438. doi: 10.1002/hbm.26399

Welton, T., Cardoso, F., Carr, J., Chan, L.-L., Deuschl, G., Jankovic, J., et al. (2021). Essential Tremor. *Nat. Rev. Dis. Prim.* 7:83. doi: 10.1038/s41572-021-00314-w