NeuroImage: Clinical 36 (2022) 103250

Contents lists available at ScienceDirect

NeuroImage: Clinical

journal homepage: www.elsevier.com/locate/ynicl

NigraNet: An automatic framework to assess nigral neuromelanin content in early Parkinson's disease using convolutional neural network

Rahul Gaurav [a,b,c,*], Romain Valabrègue [a,c], Lydia Yahia-Chérif [a,c], Graziella Mangone [a,d], Sridar Narayanan [g], Isabelle Arnulf [a,b,h], Marie Vidailhet [a,b,f], Jean-Christophe Corvol [a,d,f], Stéphane Lehéricy [a,b,c,e]

[a] Paris Brain Institute – ICM, Sorbonne University, UPMC Univ Paris 06, INSERM U1127, CNRS UMR 7225, Pitié-Salpêtrière Hospital, Paris, France
[b] Movement Investigations and Therapeutics Team (MOV'IT), ICM, Paris, France
[c] Center for NeuroImaging Research – CENIR, ICM, Paris, France
[d] INSERM, Clinical Investigation Center for Neurosciences (CIC), Pitié-Salpêtrière Hospital, Paris, France
[e] Department of Neuroradiology, Pitié-Salpêtrière Hospital, AP-HP, Paris, France
[f] Department of Neurology, Pitié-Salpêtrière Hospital, AP-HP, Paris, France
[g] McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, 3801 University Street, Montreal, Quebec, H3A 2B4, Canada
[h] Sleep Disorders Unit, Pitié-Salpêtrière Hospital, AP-HP, Paris, France

ARTICLE INFO

Keywords:
Artificial Intelligence
Deep Learning
Convolutional Neural Networks
MRI
Neuromelanin
Parkinson's Disease
Substantia Nigra

ABSTRACT

Background: Parkinson's disease (PD) demonstrates neurodegenerative changes in the substantia nigra pars compacta (SNc) using neuromelanin-sensitive (NM)-MRI. As SNc manual segmentation is prone to substantial inter-individual variability across raters, development of a robust automatic segmentation framework is necessary to facilitate nigral neuromelanin quantification. Artificial intelligence (AI) is gaining traction in the neuroimaging community for automated brain region segmentation tasks using MRI.
Objective: Developing and validating AI-based NigraNet, a fully automatic SNc segmentation framework allowing nigral neuromelanin quantification in patients with PD using NM-MRI.
Methods: We prospectively included 199 participants comprising 144 early-stage idiopathic PD patients (disease duration = 1.5 ± 1.0 years) and 55 healthy volunteers (HV) scanned using a 3 Tesla MRI including whole brain T1-weighted anatomical imaging and NM-MRI. The regions of interest (ROI) were delineated in all participants automatically using NigraNet, a modified U-net, and compared to manual segmentations performed by two experienced raters. The SNc volumes (Vol), volumes corrected by total intracranial volume (C_{vol}), normalized signal intensity (NSI) and contrast-to-noise ratio (CNR) were computed. One-way GLM-ANCOVA was performed while adjusting for age and sex as covariates. Diagnostic performance measurement was assessed using the receiver operating characteristic (ROC) analysis. Inter and intra-observer variability were estimated using Dice similarity coefficient (DSC). The agreements between methods were tested using intraclass correlation coefficient (ICC) based on a mean-rating, two-way, mixed-effects model estimates for absolute agreement. Cronbach's alpha and Bland-Altman plots were estimated to assess inter-method consistency.
Results: Using both methods, Vol, C_{vol}, NSI and CNR measurements differed between PD and HV with an effect of sex for C_{vol} and CNR. ICC values between the methods demonstrated optimal agreement for C_{vol} and CNR (ICC > 0.9) and high reproducibility (DSC: 0.80) was also obtained. The SNc measurements also showed good to excellent consistency values (Cronbach's alpha > 0.87). Bland-Altman plots of agreement demonstrated no

Abbreviations: AI, Artificial Intelligence; ANCOVA, analysis of covariance; AUC, area under the receiver operating characteristic curve; CNR, contrast-to-noise ratio; ConvNet, convolutional neural network; C_{vol}, total intracranial volume; DA, dopaminergic, DaT, dopamine transporter; DaTScan™, dopamine transporter single photon emission tomography; HV, healthy volunteers; MDS UPDRS, Movement Disorder Society Unified Parkinson's Disease Rating Scale; MoCA, Montréal cognitive assessment score; MP2RAGE, three-dimensional magnetisation-prepared two rapid gradient echo; MRI, magnetic resonance imaging; ROI, regions of interest; SD, standard deviations; SN, substantia nigra; SNc, substantia nigra pars compacta; NSI, normalized signal intensity; SD, standard deviation; TSE, turbo spin echo.

* Corresponding author at: Centre de NeuroImagerie de Recherche – CENIR, Institut du Cerveau – ICM, Hôpital Pitié-Salpêtrière, 47 Boulevard de l'Hôpital, 75013 Paris, France.
E-mail address: rahul.gaurav@icm-institute.org (R. Gaurav).

https://doi.org/10.1016/j.nicl.2022.103250
Received 23 May 2022; Received in revised form 15 October 2022; Accepted 28 October 2022
Available online 31 October 2022

association of SNc ROI measurement differences between the methods and ROI average measurements while confirming that 95 % of the data points were ranging between the limits of mean difference (d ± 1.96xSD). Percentage changes between PD and HV were −27.4 % and −17.7 % for Vol, −30.0 % and −22.2 % for C_{vol}, −15.8 % and −14.4 % for NSI, −17.1 % and −16.0 % for CNR for automatic and manual measurements respectively. Using automatic method, in the entire dataset, we obtained the areas under the ROC curve (AUC) of 0.83 for Vol, 0.85 for C_{vol}, 0.79 for NSI and 0.77 for CNR whereas in the training dataset of 0.96 for Vol, 0.95 for C_{vol}, 0.85 for NSI and 0.85 for CNR. Disease duration correlated negatively with NSI of the patients for both the automatic and manual measurements.

Conclusions: We presented an AI-based NigraNet framework that utilizes a small MRI training dataset to fully automatize the SNc segmentation procedure with an increased precision and more reproducible results. Considering the consistency, accuracy and speed of our approach, this study could be a crucial step towards the implementation of a time-saving non-rater dependent fully automatic method for studying neuromelanin changes in clinical settings and large-scale neuroimaging studies.

1. Introduction

Parkinson disease (PD) is the second-most common neurodegenerative disorder, impacting 2–3 % of the population above 65 years of age. It is characterized by the progressive loss of dopaminergic neurons in the substantia nigra pars compacta (SNc) which leads to striatal dopamine depletion (Poewe et al., 2017). The motor symptoms start to show up when the dopaminergic (DA) neuronal loss reaches a threshold of 30 % to 60 % (Cheng et al., 2010; Fearnley and Lees, 1991; Greffard et al., 2006; Hornykiewicz, 1998; Kish et al., 1988). The development of early biomarkers reflecting the degenerative process associated with the DA neuron loss could facilitate the investigation of lesion progression, relationship with clinical symptoms or screening and evaluating future disease-modifying drug candidates.

The DA neurons in the SNc contain a neuromelanin pigment that acquires paramagnetic properties when associated with metals (Zecca and Swartz, 1993). Neuromelanin-sensitive T1-weighted images acquired using different strategies are sensitive to the high paramagnetic and low magnetization transfer properties of neuromelanin (Priovoulos et al., 2018; Sulzer et al., 2018). These images show the SNc as an area of high signal intensity associated with the neuromelanin-iron compound. Previously, neuromelanin-sensitive MRI technique was validated histologically as a marker of neuromelanin (Blazejewska et al., 2013; Cassidy et al., 2019; Kitao et al., 2013; Lee et al., 2018, 2016; Sasaki et al., 2006) although more direct correlation studies between the DA neuron counts and neuromelanin signal are needed (Kitao et al., 2013). However, compared to other markers such as free water or [123I] FP-CIT (Andica et al., 2019; Burciu et al., 2017), neuromelanin may be considered as a more direct marker of DA neuron content in PD (Sulzer et al., 2018; Zecca et al., 2002).

Several studies have shown the usefulness of neuromelanin-sensitive MRI to characterize PD patients from healthy volunteers (HV) using manual or semi-automatic segmentation of SNc for computing volume and signal intensity of the SNc (Matsuura et al., 2013; Ogisu et al., 2013; Ohtsuka et al., 2013; Prasad et al., 2018a). Many concordant studies have reported reduced size and signal intensity of the SNc in PD patients using neuromelanin-sensitive imaging with a high diagnostic accuracy (Castellanos et al., 2015; Sasaki et al., 2006; Schwarz et al., 2011). However, relying on manual segmentation is a time-consuming task in large cohorts and semi-automatic methods require defining a threshold that may differ between studies. This can be challenging and prone to substantial inter-individual variability across raters and also across studies.

Currently, deep learning (DL) is used in a wide array of image processing tasks from image segmentation to extraction of various kinds of biomarkers, such as signal intensity and volume of brain structures or lesions (Burgos and Colliot, 2020; Goodfellow et al., 2016; Lecun et al., 2015). Furthermore, in DL, convolutional neural networks (ConvNet) use a simple artificial neural network architecture that has been successfully applied to image segmentation, recognition and detection of region of interest (ROI)s in medical images (Cun et al., 1998; Krizhevsky et al., 2012).

With respect to manual methods, ConvNet approaches can improve consistency, be less prone to errors and save a lot of time. While a few studies have shown the usefulness of ConvNet-based automatic SNc ROI segmentation approaches, they were tested on rather small datasets of PD patients by training the model using data augmentation techniques (Krupička et al., 2019; Le Berre et al., 2019; Shinde et al., 2019). Therefore, there is a need for a robust automatic SNc segmentation framework tested on a larger dataset of PD patients, involving fewer training data and no data augmentation that can still yield diagnostic performance similar to established ground truth, manual SNc segmentation performed by experts.

In this study, we propose a framework, NigraNet, that facilitates fully automatic SNc segmentation by employing a ConvNet-based architecture based on a modified U-net (Ronneberger et al., 2015). We aimed to investigate the neuromelanin changes in early-stage idiopathic PD patients as compared to HV, as a biomarker of SNc neurodegeneration. Thereafter, we also compared the automatic segmentation measurements to the manual measurements for the validation of our proposed framework.

2. Material and methods

2.1. Participants

In this cross-sectional, prospective, observational, case-control study, patients with early-stage idiopathic PD and HV were recruited from May 2015 to March 2020 in the ICEBERG study (ClinicalTrials.gov Identifier: NCT02305147) at the Paris Brain Institute (ICM). The PD patients were included based on the UK Brain Bank Criteria made by a movement disorder specialist (Hughes et al., 1992). They were aged between 18 and 75 years with no/minimal cognitive disturbances (Mini-Mental State Examination > 24) and a disease duration of less than 4 years. The HV did not have any current or prior history of psychiatric or neurological disorders. The local ethics committee approved this study and all participants provided written informed consent (IRB of Paris VI, RCB 2014-A00725-42).

2.2. Clinical examination

For all participants, Movement Disorders Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS) Part I for non-motor aspects of experiences of daily living; Part II for motor aspects of experiences of daily living and Part III for motor examination were measured, whereas Part IV for motor complications was obtained only for patients (Goetz et al., 2012). For computing the mean disease duration in years, we used the date of diagnosis as the starting point. Levodopa equivalent daily dose (LEDD, in mg) of patients was also collected.

2.3. MRI data acquisition

The participants were scanned using a 3 Tesla PRISMA scanner (Siemens, Erlangen, Germany) and a 64-channel receive-only head coil. The MRI protocol included whole brain three-dimensional T1-weighted (T1-w) imaging and axial turbo spin echo two-dimensional (2D) T1-w neuromelanin-sensitive imaging with a field of view restricted to the midbrain. Neuromelanin-sensitive images were carefully positioned in the axial plane perpendicular to the long axis of the brainstem and covering the locus coeruleus and the SNc. Whole brain T1-w images were acquired using a sagittal Magnetization Prepared 2 RApid Gradient Echo (MP2RAGE) with a 1-mm isovoxel size(Marques et al., 2009) and neuromelanin-sensitive images were acquired with the following parameters: with TR/TE/flip angle: 890 ms/13 ms/180°, 3 averages, voxel size: 0.4 × 0.4 × 3 mm^3, acquisition time (TA): 6:55 min.

2.4. Image analysis

Image analysis was performed in MATLAB (MathWorks Inc, MA, USA, vR2017b) using FSL (FMRIB, UK, v5.0) and Statistical Parametric Mapping (SPM, UK, v12). Image preprocessing and postprocessing were performed using MRtrix3 (UK, v3.0). The manual SNc and background ROIs were delineated using the FreeSurfer image viewer (MGH, USA, v5.3.0).

2.5. Proposed segmentation framework

The image processing pipeline for the proposed NigraNet framework comprised three major steps: (1) manual segmentation for training the network, (2) probability maps prediction using the model, and (3) postprocessing for yielding the final ROI to be used in further SNc measurements.

(i) Manual Tracing Procedure: Manual delineation of SNc contours of all the participants were performed on the neuromelanin-sensitive images by two independent trained raters (S.L. and R. G.) as the border of hyperintense area dorsal to the cerebral peduncle and ventral to the red nucleus on three slices where SNc was visible as described previously (Gaurav et al., 2021; Pyatigorskaya et al., 2017). Both raters were blind to the participant's clinical status. Moreover, in order to obtain the signal intensity ratio with the SNc ROI, a background ROI was also manually traced comprising the tegmentum and superior cerebral peduncles as shown in Fig. 1 (Gaurav et al., 2021).

(ii) Automatic Segmentation Procedure: An image-processing pipeline was developed for automatic segmentation by implementing the DL model in Python 3.6.0 using Scikit-learn (v0.21.0) and Keras library (v2.2.4) with the TensorFlow framework (v1.14) at the backend. The DL model was trained on NVIDIA-quadro P6000 (v418.67) with CUDA 10.1 and 24 GB GPU memory.

(a) Data preprocessing: All the training images were resliced to a dimension of 240 × 240 × 48. The manually segmented SNc ROIs as input features were thresholded, binarized and normalized before feeding into our training network.

(b) NigraNet architecture details: A complete schematic of the network architecture of NigraNet has been illustrated in Fig. 1.

The network consisted of an encoder in the form of contracting path and a decoder in the form of expansive path without employing any fully connected layer. The encoder part included 3 dilated three-dimensional convolution layers, each of them followed by a rectified linear unit (ReLU) activations using 3 × 3 × 3 kernel size (Nwankpa et al., 2018). We took advantage of the dilated convolutions as they help in maintaining the input resolution while attaining a larger receptive field. As a strategy to downsample the image, a three-dimensional max-pooling operation of 2 × 2 × 2 was applied to extract the maximum value. In the decoder part, the feature maps were upsampled using a learnable three-dimensional convolution transpose layer. These upsampled feature maps were concatenated with the correspondingly cropped feature map of the encoder. We employed cropping so that we do not lose the border pixels in every convolution. Thereafter, another three-dimensional convolutional layer was applied to the concatenated features that were again upsampled using a three-dimensional convolution transpose layer which were further concatenated with the correspondingly cropped feature map obtained from the first convolutional layer of the encoder. A final convolutional layer was then applied to the concatenated features using a sigmoid activation function, as this helps in predicting the probability of the SNc ROI by yielding a mask with the same dimensions and resolution as the input image.

(c) Experimentations

Evaluation metrics: Dice Similarity Coefficient (DSC) was computed as described in a previous study.(Sudre et al., 2017) To understand the loss function during the training process, Dice loss was calculated as 1-

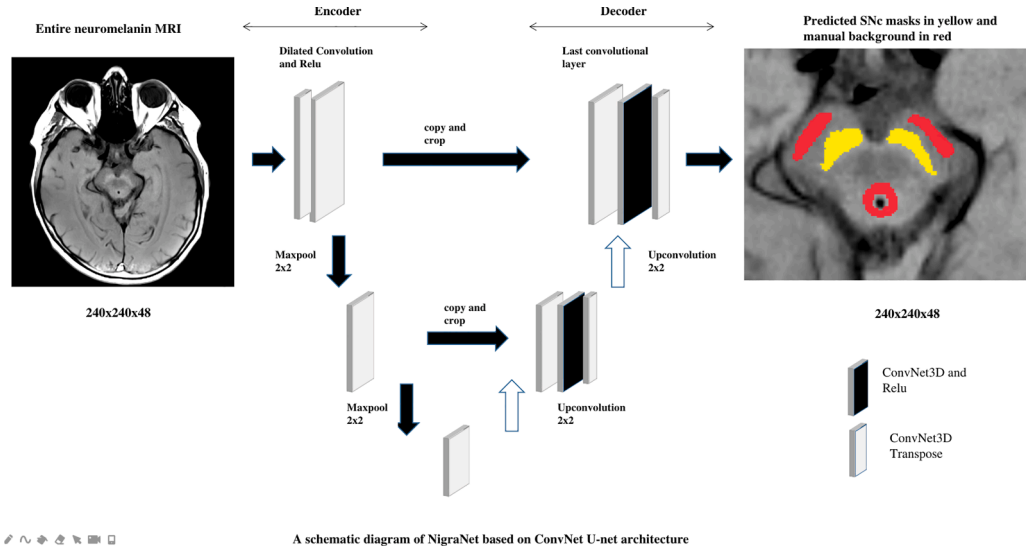


Fig. 1. NigraNet model architecture.

DSC.

Training: The model was trained using entire neuromelanin image without employing any bounding box around the targeted regions. The network was trained using the stochastic gradient-based optimizer Adam to optimize the weights associated with the connections between neurons (Kingma and Ba, 2015). We arbitrarily experimented with various hyperparameters such as epochs of 50, 100, 200, 400 and 1000; batch size from 1 to 24; validation split from 0.1 to 0.5 and shuffle of true and false before fixing the epochs to 110, shuffle to true, validation split to 0.1 along with a mini-batch size of 1. Although we experimented with data augmentation techniques by using random elastic deformation using the TorchIO library(Pérez-García et al., 2020) on the images, it did not alter our main results. Hence, from our pool of 199 manually segmented SNc ROIs of PD patients and HV, a random dataset of 60 images were randomly split using the shuffle function into 54 principal training images and 6 external validation images.

(d) Data prediction: We ran the model to predict the SNc probability mask on the remaining 139 participants which were not used for training or validation.
(e) Data postprocessing: With an aim to obtain the best segmentation performance, we binarized the predicted SNc masks by thresholding with a probability of 0.5 so that we do not have over estimation of SNc. Then we employed the maskfilter function of the MRtrix3 package(Tournier et al., 2019) in order to retain the largest connected component in the left and right SNc, yielding the final automatic ROI used in further analyses.

2.6. Quantitative analysis

The quantitative measurements were calculated on both manual and automatic SNc ROIs using our in-house algorithm written in MATLAB. (Gaurav et al., 2021) SNc volume (Vol) was obtained as the number of voxels in the neuromelanin-based ROI in the image slices multiplied by the voxel size. The total intracranial volume (TIV) was estimated as the summation of white matter, grey matter and cerebrospinal fluid volumes derived from SPM12. Thereafter, we calculated the corrected SNc volume (C_{vol}) by dividing Vol by the TIV in order to normalize for individual variations in head sizes. normalized signal intensity (NSI) and contrast-to-noise ratio (CNR) were calculated by dividing the mean signal in the SNc by the background ROI signal. In order to correct for slice-to-slice intensity variations in the 2D TSE acquisition, and for putting the SNc signal measurements on a scale that is comparable across participants and over time, we computed normalized signal intensity (NSI) and contrast-to-noise ratio (CNR) for each slice normalizing the mean signal in SNc relative to the background signal using the following formulas (Chougar et al., 2022; Gaurav et al., 2021).

$$NSI = \text{Mean\_over\_slices}\{(Sig_{SNc}/Sig_{BND})*100\}$$

$$CNR = \text{Mean\_over\_slices}\{(Sig_{SNc} - Sig_{BND})/STD_{BND}\}$$

where Sig_{SNc} is the signal intensity in SN_c ROI, Sig_{BND} the signal intensity in background ROI and STD_{BND} the standard deviation in background ROI.

2.7. Statistical analyses

Statistical analyses were performed using R (R Core Team 2019, v3.6.1) and MATLAB. We performed the between-group comparisons of demographic and clinical variables using the parametric Student's *t*-test whereas the inter-group differences in sex distribution were assessed using the Chi-square test.

Parametric tests were performed as the imaging data were normally distributed according to the Shapiro-Wilk normality test. For both automatic and manual methods, we tested the between-group difference in Vol, C_{vol}, NSI and CNR using a one-way general linear model (GLM) – analysis of covariance (ANCOVA) by considering the group (PD and HV) as the between-group factor with both age and sex as covariates.

Diagnostic performance to classify individuals as PD or HV based on SNc neuromelanin measures was assessed using receiver operating characteristic (ROC) analysis. Inter and intra-observer variability between SNc ROI segmentations were estimated using DSC. The agreements between SNc measurements derived from both segmentation methods were tested using Intraclass Correlation Coefficient (ICC) based on a mean-rating two-way mixed-effects model estimates for absolute agreement. In addition, Cronbach's alpha and Bland-Altman plots were estimated for inter-method consistency.

Thereafter, we obtained correlations based on Pearson's correlation coefficients of SNc measurements using both automatic and manual methods with clinical variables and age. An approximate multivariate permutation test was conducted while adjusting for multiple comparisons (Nichols and Holmes, 2002). Differences were considered significant at $p < 0.05$ and trends were considered between 0.05 and 0.09.

3. Results

3.1. Clinical characteristics

Altogether, we analyzed 199 participants comprising 144 patients with early-stage idiopathic PD and 55 HV. There was no significant difference in age between PD and HV, and also, no significant difference

Table 1
Demographic, clinical characteristics and Substantia Nigra Pars Compacta (SNc) measurements.

	Healthy Volunteers (HV)	Patients with Parkinson's Disease (**PD**)
Demographics		
Number of participants	55	144
Age (years)	61.0 ± 9.0	61.8 ± 9.3
Male/Female	27/28	89/55
Clinical characteristics		
Disease Duration (years)	NA	1.5 ± 1.0
MMSE score (0–30)	29.5 ± 0.8	29.0 ± 1.2 **
MDS UPDRS I (0–52)	5.0 ± 3.5	9.3 ± 3.9 **
MDS UPDRS II (0–52)	1.2 ± 1.8	8.1 ± 3.7 **
MDS UPDRS III ON (0–132)	NA	26.2 ± 7.3
MDS UPDRS III OFF (0–132)	5.6 ± 5.3	29.6 ± 7.9 **
MDS UPDRS IV (0–24)	0.0 ± 0.0	0.1 ± 0.7
Hoehn and Yahr score (0–5)	0.1 ± 0.5	2.0 ± 0.3 **
Measurements using automatic SNc ROI		
Volume (mm^3)	297.7 ± 54.8	216.0 ± 54.7 **
Corrected Volume (C_{vol})	0.20 ± 0.04	0.14 ± 0.04 **
Normalized signal intensity (NSI)	110.1 ± 1.5	108.5 ± 1.7 **
Contrast-to-Noise Ratio (CNR)	1.58 ± 0.27	1.31 ± 0.31 **
Measurements using manual SNc ROI		
Volume (mm^3)	262.4 ± 41.2	216.0 ± 55.2 **
Corrected Volume (C_{vol})	0.18 ± 0.03	0.14 ± 0.04 **
Normalized signal intensity (NSI)	111.1 ± 1.5	109.5 ± 1.6 **
Contrast-to-Noise Ratio (CNR)	1.75 ± 0.27	1.47 ± 0.34 **
Other MRI measurements		
Total Intracranial Volume (TIV, cm^3)	1465.5 ± 141.7	1519.8 ± 151.1 **

Demographical and clinical characteristics were compared using t-tests. Sex was compared using Chi-2.
SNc measurements were compared using GLM-ANCOVA while adjusting for age and sex as covariates (p value < 0.001).
*indicates significant difference between HV and PD with p value < 0.01 and ** for p value < 0.001. Abbreviations: HV: Healthy Volunteers, PD: Parkinson's Disease, MDS-UPDRS: Movement Disorders Society - Unified Parkinson's Disease Rating scale; MMSE: Mini-Mental State Examination, HY: Hoehn and Yahr scale. Data represented as mean ± standard deviation.

in sex proportion ($\chi^2 = 2.65$, $p = 0.10$) Table 1.

3.2. *Imaging results (Table 1)*

The automatic segmentation prediction took less than a second for a single participant on a modern GPU (Fig. 1).

Quantitative analysis: For both automatic and manual methods, we found a highly significant difference in all SNc measurements of Vol, C_{vol}, NSI and CNR between PD and HV ($p < 0.001$ for GLM-ANCOVA and post-hoc t-tests, Fig. 2).

For automatic and manual methods, respectively, we observed a percentage difference of −27.4 % and −17.7 % for SNc Vol, −30.0 % and –22.2 % for C_{vol}, −15.8 % and −14.4 % for NSI, −17.1 % and −16.0 % for CNR in PD compared with HV (Table 1).

ROC analyses: To discriminate PD patients from HV in the training dataset (n = 60), the ROC analyses provided areas under the curve (AUC) of 0.96 and 0.85 for Vol, 0.95 and 0.88 for C_{vol}, 0.85 and 0.85 for NSI and 0.85 and 0.85 for CNR for automatic and manual methods respectively (Fig. 3). Overall, in the entire dataset comprising both training and testing data (n = 199), AUC was 0.83 and 0.74 for Vol, 0.85 and 0.76 for C_{vol}, 0.79 and 0.78 for NSI and 0.77 and 0.76 for CNR for automatic and manual method respectively.

Effect of sex: Overall, there was a highly significant effect of sex for both C_{vol} and CNR measurements using both the automatic and manual methods ($p < 0.001$ for GLM-ANCOVA). Further separate post-hoc analysis demonstrated that both groups had significantly larger C_{vol} in females than in males ($p < 0.001$ for *t*-test). In addition, the female patients also showed significantly larger CNR than the male patients using both the methods ($p < 0.001$ for *t*-test).

3.3. *Variability of measurements*

There was high agreement between the automatic and manual segmentations (DSC: 0.80) and between the measurements performed by the two manual raters (DSC inter-observer: 0.82, DSC intra-observer: 0.85). ICC values between the methods demonstrated a strong agreement for all the SNc measurements (all the ICC values > 0.8) with optimal agreement for C_{vol} and CNR (ICC > 0.9) (Table 2).

3.4. *Reliability*

Overall, SNc ROI comparisons using volumes and signal intensities from both methods demonstrated absolute agreement ranging from good to excellent (ICC > 0.83). The consistency estimates for SNc measurements also showed good to excellent consistency values (Cronbach's alpha > 0.87).

Furthermore, visual inspection of Bland-Altman plots of agreement of SNc ROI measurements obtained from automatic and manual segmentations demonstrated no association of SNc ROI measurement differences between the methods and ROI average measurements while confirming that 95 % of the data points were ranging between the limits of mean difference (d ± 1.96xSD). (Fig. 4).

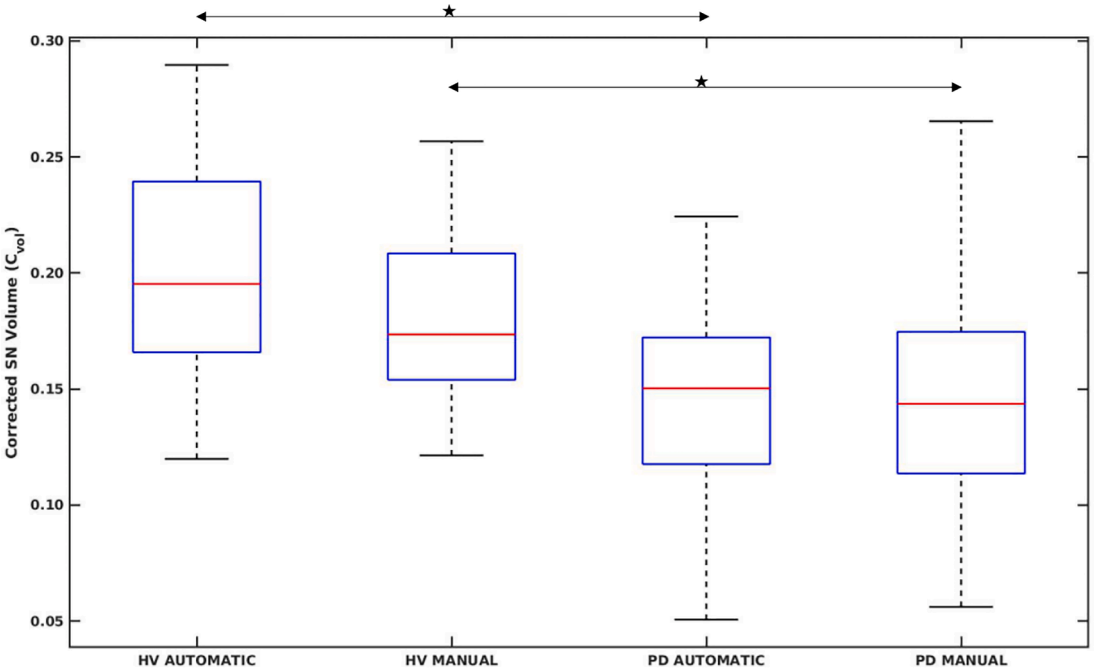
3.5. *Correlations between imaging and clinical status*

The correlations between imaging and clinical status are shown in the Table 3. Age in PD correlated negatively with Vol and positively with NSI using manual measurements but not using automatic measurements. Age in HV correlated negatively with Vol and tended to correlate negatively with C_{vol} using manual but not using automatic measurements.

The disease duration correlated negatively with NSI of the patients for both the automatic and manual measurements.

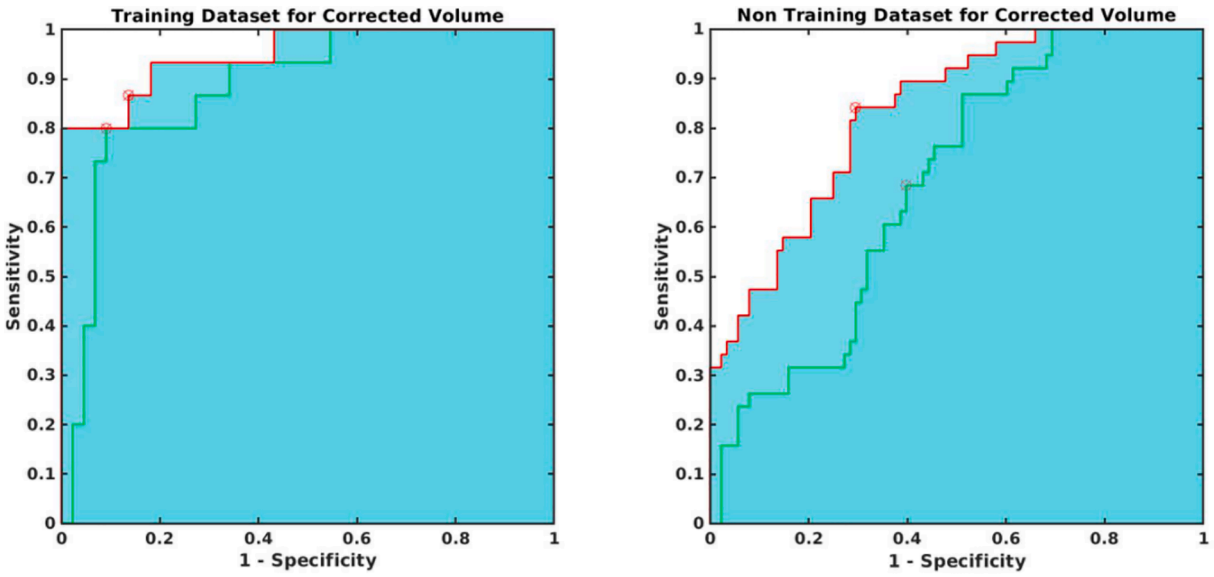
MDS UPDRS III score in OFF mode tended to correlate negatively with Vol and correlated negatively with C_{vol} in PD, and negatively with Vol and tended to correlate negatively with C_{vol} using automatic measurements in HV. No such correlations were found using manual measurements. MDS UPDRS III score in ON mode showed negative trend with Vol using automatic measurements and negative trends for both Vol and C_{vol} using manual measurements.

LEDD correlated negatively with NSI and CNR using the automatic measurements but not using the manual measurements.



Box plot of corrected SNc volume (C_{vol}) between the HV and PD using both automatic and manual segmentation methods with groups being highly significant from each other (ANOVA p value <0.001).

Fig. 2. Box plot of C_{vol} between HV and PD obtained using automatic and manual segmentation.



ROC curve based on corrected SNc volume (C_{vol}) between PD patient and HV, showing higher area under ROC value for automatic (in red) than manual (in green) segmentation method.

Fig. 3. Diagnostic performance to discriminate PD from the HV using AUC of C_{vol} using automatic method (in red) and manual method (in green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Intraclass correlation coefficients (Average raters absolute) and Cronbach coefficients between automatic and manual measurements.

Number of Judges = 3							
	ICC	**F**	**df1**	**df2**	**p**	**lower**	**upper**
Vol	0.87	7.6	199	400	1.4e-65	0.84	0.89
C_{vol}	0.90	9.6	199	400	5.7e-80	0.87	0.91
NSI	0.83	5.7	199	400	1.2e-49	0.79	0.86
CNR	0.92	13	199	400	1.4e-101	0.91	0.94
Number of Judges = 2							
Vol	0.75	3.9	199	200	5.3e-21	0.68	0.80
C_{vol}	0.80	5.0	199	200	1.9e-27	0.75	0.84
NSI	0.84	6.2	199	200	2.4e-34	0.80	0.87
CNR	0.89	9.1	199	200	8.6e-47	0.86	0.91
Cronbach coefficients							
	Alpha (3)	**Alpha (2)**					
Vol	0.87	0.75					
C_{vol}	0.90	0.81					
NSI	0.95	0.95					
CNR	0.94	0.92					

4. Discussion

This study demonstrated a DL-based robust, fast, fully automatic segmentation model of the SNc using neuromelanin-sensitive MRI. The diagnostic performance of the neuromelanin-based volume and signal intensity using the NigraNet automatic method showed comparable results with the manual method while demonstrating superior performance in C_{vol} measurements. These measurements were correlated clinically with disease duration for NSI, and LEDD for both NSI and CNR.

The NigraNet framework comprised a simple network architecture employing dilated convolutional layers that was tested on a large database of early-stage idiopathic PD patients rather than advanced-stage PD patients as in the previous studies (Le Berre et al., 2019; Shinde et al., 2019). It demonstrated reliable predictions despite training on a small number of neuromelanin-sensitive images. Taken together, our results endorsed earlier findings and demonstrated the potential of automated neuromelanin-based SNc segmentation as a powerful imaging biomarker of PD diagnosis.

4.1. Neuromelanin physiology and histological correlates

Neuromelanin is a dark pigment produced in the cytosol of catecholaminergic neurons by oxidation of excess dopamine in quinones and is subsequently stored in neuromelanin autophagic lysosome organelles where it is bound to metals such as copper, zinc and iron (Zucca et al., 2018). In PD, the synthesis of neuromelanin may be neuroprotective by removing an excess of dopamine (Zucca et al., 2014). However, beyond a certain threshold, intracellular neuromelanin accumulation may compromise neuronal function and triggers a PD-like pathology as suggested experimentally in rodents producing a tyrosinase-induced neuromelanin-like pigment in the SN (Carballo-Carbajal et al., 2019), although this needs to be confirmed in patients.

The neuromelanin-iron complex can also be found in the extracellular compartment after neuronal death (Beach et al., 2007). The extracellular neuromelanin-iron complex may then participate to the neurotoxicity by neuroinflammation processes and activated microglia (Zhang et al., 2011).

The neuromelanin-iron complex is paramagnetic (Zecca and Swartz, 1993) and thus hyperintense on T1-weighted images at 3 Tesla, using gradient echo with magnetization transfer,(Chen et al., 2014; Priovoulos et al., 2018) black blood delay alternating with nutation for tailored excitation-prepared T1-weighted variable flip angle turbo spin echo (DANTE T1-SPACE),(Oshima et al., 2020) spin echo (Chen et al., 2014; Kitao et al., 2013; Sasaki et al., 2006; Schwarz et al., 2011) and also at 7 Tesla (Blazejewska et al., 2013).

4.2. Neurodegeneration pattern in SNc

Neuromelanin-based SNc volume measurements showed variable results (Castellanos et al., 2015; Isaias et al., 2016; Pyatigorskaya et al., 2018; Takahashi et al., 2018) depending on several factors such as the calculation technique, the number of segmented slices, the type of sequence and other imaging parameters like slice thickness that often leads to partial voluming. We obtained larger segmentation with a

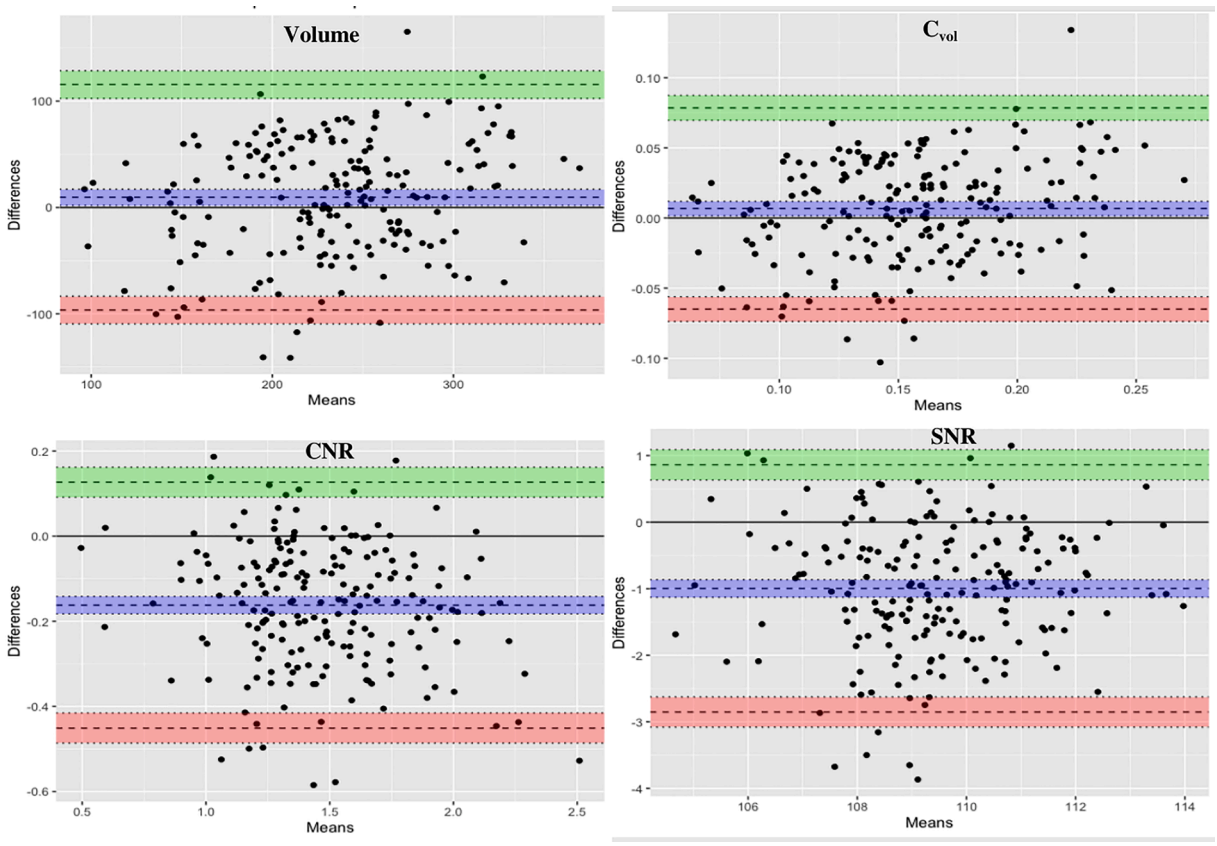


Fig. 4. Bland Altman plots.

higher SNc Vol values in the automated method as compared to the manual method. This could be due to several potential reasons such as the fact that we resized the ROIs and raw images to 240 × 240 × 48, yielding around 10–15 slices after prediction.

We observed highly significant differences in volume and signal intensity in early PD. Neuromelanin-based SNc volume in PD for automatic and manual methods decreased by 27.4 % and 17.7 % respectively whereas NSI decreased by 16.9 % and 14.9 % respectively as compared to the HV. Previously,(Gaurav et al., 2021) we showed lower Vol reduction of around 11 % possibly because we used manual segmentations on raw neuromelanin images on 3 slices involving no image resizing as in this current study. Some studies have demonstrated similar reductions in size of 22 % to 29 % in de novo patients (Reimão et al., 2016, 2015; Wang et al., 2019, 2018) and some in the range of 29 % to 47 % either using measurements based on SN width or area. (Isaias et al., 2016; Reimão et al., 2016; Schwarz et al., 2011) These reductions varied across studies possibly because of patient selection and the methodology used. Recent literature also demonstrated variations in signal intensity reductions such as −21 % in de novo PD(Wang et al., 2018) to −48 % in advance stage PD (Ohtsuka et al., 2013). However, some studies reported greater signal intensity variations in the lateral part of this structure (Biondetti et al., 2020; Martín-Bastida et al., 2019; Ohtsuka et al., 2014; Prasad et al., 2018b; Schwarz et al., 2017; Takahashi et al., 2019; Wang et al., 2019, 2018).

4.3. AI-based models for MRI segmentations

Previous studies using DL-based SNc segmentation included relatively smaller datasets of 36 PD patients,(Krupička et al., 2019) 45 PD patients (Le Berre et al., 2019) or 60 PD patients (Shinde et al., 2019). This study using ConvNet(Le Berre et al., 2019) obtained a percentage of Vol reduction in PD compared with HV of 46.4 % using automatic method and 42.8 % using the manual method, and AUC of 0.95 for both automatic and manual measurements. In our present study, PD patients had a lower percentage Vol decrease and AUC relative to HV than the aforementioned studies, likely because we examined early PD patients.

Studies also reported better performance of SNc segmentation using a U-net model compared to threshold segmentation(Krupička et al., 2019; Le Berre et al., 2019) for discriminating patients from HV, particularly in the superior segmented SNc slices, with a Vol reduction of 26.2 % between HV and PD for all slices, which was in line with ours (Krupička et al., 2019). More studies utilizing inferior, middle and superior SNc slices separately are warranted to confirm these results (Krupička et al., 2019).

The study (Le Berre et al., 2019) examined patients with more advanced PD with disease duration of 10.0 ± 6.2 years in the principal dataset and 15.8 ± 13.3 years in the external dataset. Another study (Shinde et al., 2019) included patients with moderately advanced PD with disease duration of 6.3 ± 4.1 years where ConvNet showed better performance than the radiomics based classifiers. Although this study (Shinde et al., 2019) trained the model using a boxed region around the brain stem on the axial slices of the neuromelanin images.

Our model had the advantage of not using a boxed region around the SNc and rather employing the entire neuromelanin image as the training input to minimize preprocessing steps, and the model was tested on early-stage idiopathic PD patients of disease duration of 1.5 ± 1.0 years.

We attempted to improve the accuracy of the model using data augmentation by applying random elastic deformation. However, this did not alter our segmentation accuracy, as possibly it needs further experimentation using random motion, gamma, blur or flip (Pérez-García et al., 2020). Recently, we also used a similar approach on patients with isolated REM sleep behavior disorder, (Gaurav et al., 2022) the prodromal phase of parkinsonism (Iranzo et al., 2014; Pyatigorskaya et al., 2017). Lastly, a recent study in PD patients employed U-net to automatically segment the locus coeruleus, which can also be visualized using the same neuromelanin-sensitive MRI acquisitions (Dünnwald

Table 3
Correlations of SNc measurements with age and clinical scores for automatic and manual measurements.

	AUTOMATIC METHOD															
	Healthy Volunteers (HV)								Patients with Parkinson's Disease (PD)							
	Volume		C_{vol}		NSI		CNR		Volume		C_{vol}		NSI		CNR	
	r	p	r	p	r	p	r	p	r	p	r	p	r	p	r	p
AGE	−0.111	0.206	−0.058	0.346	0.125	0.189	0.117	0.205	0.039	0.323	0.058	0.248	0.050	0.277	0.054	0.244
MDS UPDRS III OFF	−0.327	**0.007**	−0.220	*0.057*	0.319	**0.009**	0.112	0.213	−0.127	*0.062*	−0.139	**0.048**	−0.015	0.438	−0.027	0.376
MMSE	0.079	0.273	−0.020	0.458	−0.065	0.329	−0.041	0.392	0.143	*0.051*	0.121	*0.075*	−0.059	0.241	−0.040	0.319
Schwab and England	0.125	0.188	0.003	0.485	−0.189	*0.098*	−0.060	0.340	0.065	0.217	0.037	0.330	0.080	0.175	0.057	0.256
HY Score	−0.197	*0.082*	−0.073	0.319	0.228	*0.059*	0.031	0.417	0.116	*0.085*	0.122	*0.069*	0.013	0.440	−0.002	0.493
MDS UPDRS IV	NA	NA	NA	NA	NA	NA	NA	NA	0.006	0.512	0.048	0.304	−0.084	0.153	−0.100	*0.098*
MDS UPDRS III ON	NA	NA	NA	NA	NA	NA	NA	NA	−0.120	*0.093*	−0.112	0.108	0.009	0.465	−0.011	0.455
DISEASE DURATION(years)	NA	NA	NA	NA	NA	NA	NA	NA	−0.011	0.446	−0.133	*0.056*	−0.261	**0.001**	−0.046	0.279
LEDD	NA	NA	NA	NA	NA	NA	NA	NA	0.055	0.265	0.013	0.439	−0.152	**0.029**	−0.157	**0.026**
	MANUAL METHOD															
	Healthy Volunteers (HV)								Parkinson's Disease (PD)							
	Volume		C_{vol}		NSI		CNR		Volume		C_{vol}		NSI		CNR	
	r	p	r	p	r	p	r	p	r	p	r	p	r	p	r	p
AGE	−0.302	**0.016**	−0.222	*0.051*	0.069	0.305	0.109	0.233	−0.150	**0.040**	−0.115	*0.090*	0.145	**0.042**	0.118	*0.073*
MDS UPDRS III OFF	−0.083	0.270	0.024	0.432	−0.010	0.470	−0.118	0.205	−0.066	0.222	−0.088	0.146	−0.076	0.182	−0.077	0.176
MMSE	−0.087	0.264	−0.203	*0.074*	0.031	0.412	0.056	0.326	−0.041	0.327	−0.046	0.299	−0.005	0.473	0.016	0.427
Schwab and England	0.005	0.491	−0.141	0.160	−0.090	0.257	0.037	0.398	−0.019	0.417	−0.038	0.310	0.108	*0.098*	0.089	0.153
HY	0.012	0.458	0.166	0.117	−0.313	**0.024**	−0.239	**0.038**	0.155	**0.034**	0.155	**0.032**	−0.042	0.296	−0.049	0.275
MDS UPDRS IV	NA	NA	NA	NA	NA	NA	NA	NA	0.062	0.240	0.106	*0.096*	−0.174	**0.027**	−0.150	**0.031**
MDS UPDRS III ON	NA	NA	NA	NA	NA	NA	NA	NA	−0.139	*0.059*	−0.133	*0.069*	0.004	0.483	−0.014	0.436
DISEASE DURATION (years)	NA	NA	NA	NA	NA	NA	NA	NA	−0.051	0.270	−0.108	0.102	−0.246	**0.001**	−0.082	0.161
LEDD	NA	NA	NA	NA	NA	NA	NA	NA	−0.074	0.197	−0.096	0.128	−0.092	0.137	−0.099	0.105

et al., 2021).

Taken together, our framework not only needed few training images but did not need any data augmentation to yield precise SNc segmentations. Furthermore, our untrained testing data sample size was more than double the size of the training data, suggesting the robustness of our model.

4.4. Effect of clinical parameters, sex and age

Overall, the correlations between the signal intensity changes and disease duration changes using both methods are in line with results of previous studies (Isaias et al., 2016; Okuzumi et al., 2019; Prasad et al., 2018b; Taniguchi et al., 2018). Correlations results for age, MDS-UPDRS III scores, and LEDD differed between the two methods. Using the manual method, age correlated with NSI and Vol in PD in line with our previous study (Gaurav et al., 2021) and also with Vol in HV. Using the automatic method, we did not observe any correlation between neuromelanin SNc measurements in HV and age. Although this dissimilarity in correlations should be interpreted with caution, this could be due to the hypothesis in line with a previous study which demonstrated a plateau of neuromelanin in midlife including the 5th-6th decades (Xing et al., 2018). MDS-UPDRS III scores correlated with C_{vol} in PD, in line with previous studies (Gaurav et al., 2021; Okuzumi et al., 2019; Schwarz et al., 2017; Taniguchi et al., 2018; Xing et al., 2022) and Vol in HV using the automatic but not the manual method. LEDD showed a significantly negative correlation with NSI and CNR using the automatic method. Although this correlation was expected as LEDD increases with disease duration and dopaminergic denervation, it was not found using the manual method, in line with our previous study (Gaurav et al., 2021). Overall, the most reproducible correlations were therefore observed between NSI and disease duration. The other correlations were more variable between methods and between different studies and therefore probably less robust. There could be several reasons to this such as larger segmentation in the automatic method compared to the manual method. Also, image resampling procedure in this study could also be a potential reason. For instance, automatic SNc predictions often yielded more slices than manual images, as the model predicted SNc mask wherever they were visible in the image slices. Thus, this difference highlights the fact that although both methods provided overall similar results, yet they did not perform identical SNc segmentation.

We observed a highly significant effect of sex for both C_{vol} and CNR measurements with higher C_{vol} in female participants and higher CNR in female patients using both segmentation methods, in line with our previous study (Gaurav et al., 2021). The sex differences in signal intensity were not observed in HV with a much larger age range (Xing et al., 2018). The DAT studies in HV have also reported higher striatal binding in women compared with men (Kaasinen et al., 2015; Muellner et al., 2015; Varrone et al., 2013).

4.5. Limitations and challenges

Although our aforementioned study showed promising results, it has several limitations and we intend to improve this framework in various ways in the future. Firstly, there is a need for validation using external cohorts, preferably with different image acquisition parameters or in clinical settings. Secondly, we used thick 2D slices (2- to 3-mm) which are prone to partial voluming; the implementation of three-dimensional acquisitions may possibly improve the results (Oshima et al., 2020, 2018). Thirdly, other acquisition schemes like magnetization transfer can possibly enhance the signal intensity (Langley et al., 2016; Liu et al., 2020; Schwarz et al., 2011). Fourthly, our method did not allow studying the topography of changes as done in previous studies that segmented the SNc into ventrolateral and dorsomedial compartments (Xing et al., 2022) or using voxel wise method (Biondetti et al., 2020; Cassidy et al., 2019). Lastly, an important question would be to understand what the effect of a drug might look like. We can hypothesize that

a drug acting on the neuromelanin content or its paramagnetic effect would modify the signal intensity, while a drug slowing neuronal death would slow the loss of SNc volume as well.

5. Conclusion

We presented a framework NigraNet utilizing small MRI training dataset to fully automatize the SNc segmentation procedure by eliminating the human factor as much as possible in order to process more data in a rather shorter time with an increased precision and more reproducible results. We showed comparable performance with ground truth for detecting neuromelanin-based volume and signal changes in the SNc. Considering the consistency, accuracy and speed of our approach, this study could be a crucial step towards the implementation of a time-saving, non-rater-dependent, fully automatic method for studying neuromelanin changes in clinical settings and large-scale neuroimaging studies.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data can be made available from the corresponding author on reasonable request

Acknowledgements

The authors would like to thank Energipole (M. Mallart), M.Villain and Société Française de Médecine Esthétique (M. Legrand) for unrestricted support for Research on Parkinson's disease. We would also like to thank all of the participants involved in the ICEBERG study.

Author Contributions

1. Conceived and designed the study, contributed analysis tools; performed the analyses; wrote the manuscript: Rahul Gaurav
2. Collected the data, reviewed the manuscript: Romain Valabrègue, Lydia Yahia-Chérif, Graziella Mangone, Isabelle Arnulf, Marie Vidailhet and Jean-Christophe Corvol
3. Conceived and designed the study; reviewed the manuscript: Stéphane Lehéricy

Financial Disclosure/Conflict of interest concerning the research related to the manuscript

Rahul Gaurav and Stéphane Lehéricy received grant funding from Biogen Inc. United States of America.

Romain Valabrègue, Lydia Yahia-Chérif, Graziella Mangone and Marie Vidailhet have nothing to report.

Sridar Narayanan received research funding from the Canadian Institutes of Health Research, the International Progressive MS Alliance, the Myelin Repair Foundation and Immunotec. He also received a speaker's honorarium from Novartis Canada, and personal compensation from NeuroRx Research.

Jean-Christophe Corvol has served in advisory boards for Air Liquide, Biogen Inc., Denali, Ever Pharma, Idorsia, Prevail Therapeutic, Theranexus, UCB; has received grants from Sanofi and the Michael J Fox Foundation.

Isabelle Arnulf received honoraria from Idorsia Pharma and speaking engagement from UCB Pharma, unrelated to this study.

Funding sources

This work was supported by grants from DHOS-Inserm, France Parkinson, Ecole des NeuroSciences de Paris (ENP), Fondation pour la Recherche Médicale (FRM), and the Investissements d'Avenir, IAIHU-06 (Paris Institute of Neurosciences – IHU), ANR-11-INBS-0006, Fondation d'Entreprise EDF, BIOGEN Inc., Fondation Thérèse and René Planiol, Fondation Saint-Michel, Unrestricted support for Research on Parkinson's disease from Energipole (M. Mallart) and Société Française de Médecine Esthétique (M Legrand).

References

Andica, C., Kamagata, K., Hatano, T., Saito, A., Uchida, W., Ogawa, T., Takeshige-Amano, H., Zalesky, A., Wada, A., Suzuki, M., Hagiwara, A., Irie, R., Hori, M., Kumamaru, K.K., Oyama, G., Shimo, Y., Umemura, A., Pantelis, C., Hattori, N., Aoki, S., 2019. Free-Water imaging in white and gray matter in Parkinson's Disease. Cells 8, 839. https://doi.org/10.3390/cells8080839.

Beach, T.G., Sue, L.I., Walker, D.G., Lue, L.F., Connor, D.J., Caviness, J.N., Sabbagh, M. N., Adler, C.H., 2007. Marked microglial reaction in normal aging human substantia nigra: correlation with extraneuronal neuromelanin pigment deposits. Acta Neuropathol. 114, 419–424. https://doi.org/10.1007/s00401-007-0250-5.

Biondetti, E., Gaurav, R., Yahia-cherif, L., Mangone, G., Pyatigorskaya, N., Valabre, R., Ewenczyk, C., Hutchison, M., Franc, C., Corvol, J., Vidailhet, M., 2020. Spatiotemporal changes in substantia nigra neuromelanin content in Parkinson's disease. Brain. https://doi.org/10.1093/brain/awaa216.

Blazejewska, A.I., Schwarz, S.T., Pitiot, A., Stephenson, M.C., Lowe, J., Bajaj, N., Bowtell, R.W., Auer, D.P., Gowland, P.A., 2013. Visualization of nigrosome 1 and its loss in PD: Pathoanatomical correlation and in vivo 7 T MRI. Neurology 81, 534–540. https://doi.org/10.1212/WNL.0b013e31829e6fd2.

Burciu, R.G., Ofori, E., Archer, D.B., Wu, S.S., Pasternak, O., McFarland, N.R., Okun, M. S., Vaillancourt, D.E., 2017. Progression marker of Parkinson's disease: A 4-year multi-site imaging study. Brain 140, 2183–2192. https://doi.org/10.1093/brain/awx146.

Burgos, N., Colliot, O., 2020. Machine learning for classification and prediction of brain diseases: recent advances and upcoming challenges. Curr. Opin. Neurol. 33, 439–450. https://doi.org/10.1097/WCO.0000000000000838.

Carballo-Carbajal, I., Laguna, A., Romero-Giménez, J., Cuadros, T., Bové, J., Martinez-Vicente, M., Parent, A., Gonzalez-Sepulveda, M., Peñuelas, N., Torra, A., Rodríguez-Galván, B., Ballabio, A., Hasegawa, T., Bortolozzi, A., Gelpi, E., Vila, M., 2019. Brain tyrosinase overexpression implicates age-dependent neuromelanin production in Parkinson's disease pathogenesis. Nat. Commun. 10 https://doi.org/10.1038/s41467-019-08858-y.

Cassidy, C.M., Zucca, F.A., Girgis, R.R., Baker, S.C., Weinstein, J.J., Sharp, M.E., Bellei, C., Valmadre, A., Vanegas, N., Kegeles, L.S., Brucato, G., Kang, U.J., Sulzer, D., Zecca, L., Abi-Dargham, A., Horga, G., 2019. Neuromelanin-sensitive MRI as a noninvasive proxy measure of dopamine function in the human brain. Proc. Natl. Acad. Sci. U. S. A. 116, 5108–5117. https://doi.org/10.1073/pnas.1807983116.

Castellanos, G., Fernández-Seara, M.A., Lorenzo-Betancor, O., Ortega-Cubero, S., Puigvert, M., Uranga, J., Vidorreta, M., Irigoyen, J., Lorenzo, E., Muñoz-Barrutia, A., Ortiz-de-Solorzano, C., Pastor, P., Pastor, M.A., 2015. Automated neuromelanin imaging as a diagnostic biomarker for Parkinson's Disease. Mov. Disord. 30, 945–952. https://doi.org/10.1002/mds.26201.

Chen, X., Huddleston, D.E., Langley, J., Ahn, S., Barnum, C.J., Factor, S.A., Levey, A.I., Hu, X., 2014. Simultaneous imaging of locus coeruleus and substantia nigra with a quantitative neuromelanin MRI approach. Magn. Reson. Imaging 32, 1301–1306. https://doi.org/10.1016/j.mri.2014.07.003.

Cheng, H.C., Ulane, C.M., Burke, R.E., 2010. Clinical progression in Parkinson disease and the neurobiology of axons. Ann. Neurol. 67, 715–725. https://doi.org/10.1002/ana.21995.

Chougar, L., Arsovic, E., Gaurav, R., Biondetti, E., Faucher, A., Valabrègue, R., Pyatigorskaya, N., Dupont, G., Lejeune, F., Cormier, F., Corvol, J., Vidailhet, M., Degos, B., Grabli, D., Lehéricy, S., 2022. Regional selectivity of neuromelanin changes in the substantia Nigra in Atypical Parkinsonism. Mov. Disord. 1–12 https://doi.org/10.1002/mds.28988.

LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W., Jackel, L., 1998. LeNet-5: handwritten digit recognition with a back-propagation network. Adv. Neural Inf. Process. Syst. 39, 149. https://doi.org/10.1111/dsu.12130.

Dünnwald, M., Ernst, P., Düzel, E., Tönnies, K., Betts, M.J., Oeltze-Jafra, S., 2021. Fully automated deep learning-based localization and segmentation of the locus coeruleus in aging and Parkinson's disease using neuromelanin-sensitive MRI. Int. J. Comput. Assist. Radiol. Surg. 16, 2129–2135. https://doi.org/10.1007/s11548-021-02528-5.

Fearnley, J.M., Lees, A.J., 1991. Ageing and parkinson's disease: Substantia nigra regional selectivity. Brain 114, 2283–2301. https://doi.org/10.1093/brain/114.5.2283.

Gaurav, R., Yahia-Cherif, L., Pyatigorskaya, N., Mangone, G., Biondetti, E., Valabrègue, R., Ewenczyk, C., Hutchison, R.M., Cedarbaum, J.M., Corvol, J.C., Vidailhet, M., Lehéricy, S., 2021. Longitudinal changes in neuromelanin MRI signal in Parkinson's disease: a progression marker. Mov. Disord. 36, 1592–1602. https://doi.org/10.1002/mds.28531.

Gaurav, R., Pyatigorskaya, N., Biondetti, E., Valabrègue, R., Yahia-Cherif, L., Mangone, G., Leu-Semenescu, S., Corvol, J., Vidailhet, M., Arnulf, I., Lehéricy, S., 2022. Deep learning-based neuromelanin MRI changes of isolated REM sleep behavior disorder. Mov. Disord. 1–7 https://doi.org/10.1002/mds.28933.

Goetz, C.G., Stebbins, G.T., Tilley, B.C., 2012. Calibration of unified Parkinson's disease rating scale scores to Movement Disorder Society-unified Parkinson's disease rating scale scores. Mov. Disord. 27, 1239–1242. https://doi.org/10.1002/mds.25122.

Goodfellow, I., Bengio, Y., Courville, A., Bengio, Y., 2016. Deep Learning. The MIT Press.

Greffard, S., Verny, M., Bonnet, A.M., Beinis, J.Y., Gallinari, C., Meaume, S., Piette, F., Hauw, J.J., Duyckaerts, C., 2006. Motor score of the unified Parkinson disease rating scale as a good predictor of lewy body-associated neuronal loss in the substantia nigra. Arch. Neurol. 63, 584–588. https://doi.org/10.1001/archneur.63.4.584.

Hornykiewicz, O., 1998. Biochemical aspects of Parkinson's disease. Neurology 51, S2–S9. https://doi.org/10.1212/wnl.51.2_suppl_2.s2.

Hughes, A.J., Daniel, S.E., Kilford, L., Lees, A.J., 1992. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. J. Neurol. Neurosurg. Psychiatry 55, 181–184. https://doi.org/10.1136/jnnp.55.3.181.

Iranzo, A., Fernández-Arcos, A., Tolosa, E., Serradell, M., Molinuevo, J.L., Valldeoriola, F., Gelpi, E., Vilaseca, I., Sánchez-Valle, R., Lladó, A., Gaig, C., Santamaría, J., 2014. Neurodegenerative disorder risk in idiopathic REM sleep behavior disorder: Study in 174 patients. PLoS One 9. https://doi.org/10.1371/journal.pone.0089741.

Isaias, I.U., Trujillo, P., Summers, P., Marotta, G., Mainardi, L., Pezzoli, G., Zecca, L., Costa, A., 2016. Neuromelanin imaging and dopaminergic loss in parkinson's disease. Front. Aging Neurosci. 8, 196. https://doi.org/10.3389/fnagi.2016.00196.

Kaasinen, V., Joutsa, J., Noponen, T., Johansson, J., Seppänen, M., 2015. Effects of aging and gender on striatal and extrastriatal [123I]FP-CIT binding in Parkinson's disease. Neurobiol. Aging 36, 1757–1763. https://doi.org/10.1016/j.neurobiolaging.2015.01.016.

Kingma, D.P., Ba, J.L., 2015. Adam: A method for stochastic optimization. 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc. 1–15.

Kish, S.J., Shannak, K., Hornykiewicz, O., 1988. Uneven pattern of dopamine loss in the striatum of patients with idiopathic Parkinson's disease. N. Engl. J. Med. 318, 876–880. https://doi.org/10.1056/NEJM198804073181402.

Kitao, S., Matsusue, E., Fujii, S., Miyoshi, F., Kaminou, T., Kato, S., Ito, H., Ogawa, T., 2013. Correlation between pathology and neuromelanin MR imaging in Parkinson's disease and dementia with Lewy bodies. Neuroradiology 55, 947–953. https://doi.org/10.1007/s00234-013-1199-9.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet classification with deep convolutional neural networks. Commun. ACM 60, 84–90. https://doi.org/10.1145/3065386.

Krupička, R., Mareček, S., Malá, C., Lang, M., Klempíř, O., Duspivová, T., Široká, R., Jarošíková, T., Keller, J., Šonka, K., Růžička, E., Dušek, P., 2019. Automatic substantia nigra segmentation in neuromelanin-sensitive MRI by deep neural network in patients with prodromal and manifest synucleinopathy. Physiol. Res. 68, S453–S458. https://doi.org/10.33549/PHYSIOLRES.934380.

Langley, J., Huddleston, D.E., Merritt, M., Chen, X., McMurray, R., Silver, M., Factor, S. A., Hu, X., 2016. Diffusion tensor imaging of the substantia nigra in Parkinson's disease revisited. Hum. Brain Mapp. 37, 2547–2556. https://doi.org/10.1002/hbm.23192.

Le Berre, A., Kamagata, K., Otsuka, Y., Andica, C., Hatano, T., Saccenti, L., Ogawa, T., Takeshige-Amano, H., Wada, A., Suzuki, M., Hagiwara, A., Irie, R., Hori, M., Oyama, G., Shimo, Y., Umemura, A., Hattori, N., Aoki, S., 2019. Convolutional neural network-based segmentation can help in assessing the substantia nigra in neuromelanin MRI. Neuroradiology 61, 1387–1395. https://doi.org/10.1007/s00234-019-02279-w.

Lecun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539.

Lee, J.H., Baek, S.Y., Song, Y.K., Lim, S., Lee, H., Nguyen, M.P., Kim, E.J., Huh, G.Y., Chun, S.Y., Cho, H.J., 2016. The Neuromelanin-related T2* Contrast in Postmortem Human Substantia Nigra with 7T MRI. Sci. Rep. 6, 32647. https://doi.org/10.1038/srep32647.

Lee, H., Baek, S.Y., Chun, S.Y., Lee, J.H., Cho, H.J., 2018. Specific visualization of neuromelanin-iron complex and ferric iron in the human post-mortem substantia nigra using MR relaxometry at 7T. Neuroimage 172, 874–885. https://doi.org/10.1016/j.neuroimage.2017.11.035.

Liu, M., Li, F., Yan, H., Wang, K., Ma, Y., Shen, L., Xu, M., 2020. A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease. Neuroimage 208. https://doi.org/10.1016/j.neuroimage.2019.116459.

Marques, J., Kober, T., Krueger, G., van der Zwaag, W., van de Moortele, P., Gruetter, R., 2009. MP2RAGE contrast optimization at 7T and applications. In: Proceedings 17th Scientific Meeting, International Society for Magnetic Resonance in Medicine. p. 2698.

Martín-Bastida, A., Lao-Kaim, N.P., Roussakis, A.A., Searle, G.E., Xing, Y., Gunn, R.N., Schwarz, S.T., Barker, R.A., Auer, D.P., Piccini, P., 2019. Relationship between neuromelanin and dopamine terminals within the Parkinson's nigrostriatal system. Brain 142, 2023–2036. https://doi.org/10.1093/brain/awz120.

Matsuura, K., Maeda, M., Yata, K., Ichiba, Y., Yamaguchi, T., Kanamaru, K., Tomimoto, H., 2013. Neuromelanin magnetic resonance imaging in Parkinson's disease and multiple system atrophy. Eur. Neurol. 70, 70–77. https://doi.org/10.1159/000350291.

Muellner, J., Gharrad, I., Habert, M.O., Kas, A., Martini, J.B., Cormier-Dequaire, F., Tahiri, K., Vidailhet, M., Meier, N., Brice, A., Schuepbach, M., Mallet, A., Hartmann, A., Corvol, J.C., 2015. Dopaminergic denervation severity depends on COMT Val158Met polymorphism in Parkinson's disease. Park. Relat. Disord. 21, 471–476. https://doi.org/10.1016/j.parkreldis.2015.02.009.

Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: A primer with examples. Hum. Brain Mapp. 15, 1–25. https://doi.org/10.1002/hbm.1058.

Nwankpa, C., Ijomah, W., Gachagan, A., Marshall, S., 2018. Activation Functions: Comparison of trends in Practice and Research for Deep Learning 1–20.

Ogisu, K., Kudo, K., Sasaki, M., Sakushima, K., Yabe, I., Sasaki, H., Terae, S., Nakanishi, M., Shirato, H., 2013. 3D neuromelanin-sensitive magnetic resonance imaging with semi-automated volume measurement of the substantia nigra pars compacta for diagnosis of Parkinson's disease. Neuroradiology 55, 719–724. https://doi.org/10.1007/s00234-013-1171-8.

Ohtsuka, C., Sasaki, M., Konno, K., Koide, M., Kato, K., Takahashi, J., Takahashi, S., Kudo, K., Yamashita, F., Terayama, Y., 2013. Changes in substantia nigra and locus coeruleus in patients with early-stage Parkinson's disease using neuromelanin-sensitive MR imaging. Neurosci. Lett. 541, 93–98. https://doi.org/10.1016/j.neulet.2013.02.012.

Ohtsuka, C., Sasaki, M., Konno, K., Kato, K., Takahashi, J., Yamashita, F., Terayama, Y., 2014. Differentiation of early-stage parkinsonisms using neuromelanin-sensitive magnetic resonance imaging. Park. Relat. Disord. 20, 755–760. https://doi.org/10.1016/j.parkreldis.2014.04.005.

Okuzumi, A., Hatano, T., Kamagata, K., Hori, M., Mori, A., Oji, Y., Taniguchi, D., Daida, K., Shimo, Y., Yanagisawa, N., Nojiri, S., Aoki, S., Hattori, N., 2019. Neuromelanin or DaT-SPECT: which is the better marker for discriminating advanced Parkinson's disease? Eur. J. Neurol. 26, 1408–1416. https://doi.org/10.1111/ene.14009.

Oshima, S., Fushimi, Y., Okada, Tomohisa, Hinoda, T., Yamamoto, T., Fukutomi, H., Yamamoto, A., Okada, Tsutomu, Grinstead, J., Ahn, S., Togashi, K., T-space, D., 2018. Visualization of Substantia Nigra Pars Compacta: MPRAGE vs . DANTE T1-SPACE, in: Joint Annual Meeting at International Society for Magnetic Resonance in Medicine & European Society for Magnetic Resonance in Medicine and Biology (ISMRM- ESMRMB2018), (France).

Oshima, S., Fushimi, Y., Okada, T., Nakajima, S., Yokota, Y., Shima, A., Grinstead, J., Ahn, S., Sawamoto, N., Takahashi, R., Nakamoto, Y., 2020. Neuromelanin-sensitive magnetic resonance imaging using DANTE pulse. Mov. Disord. 1–10 https://doi.org/10.1002/mds.28417.

Pérez-García, F., Sparks, R., Ourselin, S., 2020. TorchIO: a Python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning.

Poewe, W., Seppi, K., Tanner, C.M., Halliday, G.M., Brundin, P., Volkmann, J., Schrag, A. E., Lang, A.E., 2017. Parkinson disease. Nat. Rev. Dis. Prim. 3, 1–21. https://doi.org/10.1038/nrdp.2017.13.

Prasad, S., Saini, J., Yadav, R., Pal, P.K., 2018a. Motor asymmetry and neuromelanin imaging: concordance in Parkinson's disease. Park. Relat. Disord. 53, 28–32. https://doi.org/10.1016/j.parkreldis.2018.04.022.

Prasad, S., Stezin, A., Lenka, A., George, L., Saini, J., Yadav, R., Pal, P.K., 2018b. Three-dimensional neuromelanin-sensitive magnetic resonance imaging of the substantia nigra in Parkinson's disease. Eur. J. Neurol. 25, 680–686. https://doi.org/10.1111/ene.13573.

Priovoulos, N., Jacobs, H.I.L., Ivanov, D., Uludağ, K., Verhey, F.R.J., Poser, B.A., 2018. High-resolution in vivo imaging of human locus coeruleus by magnetization transfer MRI at 3T and 7T. Neuroimage 168, 427–436. https://doi.org/10.1016/j.neuroimage.2017.07.045.

Pyatigorskaya, N., Gaurav, R., Arnaldi, D., Leu-Semenescu, S., Yahia-Cherif, L., Valabregue, R., Vidailhet, M., Arnulf, I., Lehéricy, S., 2017. Magnetic Resonance Imaging Biomarkers to Assess Substantia Nigra Damage in Idiopathic Rapid Eye Movement Sleep Behavior Disorder. Sleep 40, 1–8. https://doi.org/10.1093/sleep/zsx149.

Pyatigorskaya, N., Magnin, B., Mongin, M., Yahia-Cherif, L., Valabregue, R., Arnaldi, D., Ewenczyk, C., Poupon, C., Vidailhet, M., Lehéricy, X.S., 2018. Comparative study of MRI biomarkers in the substantia nigra to discriminate idiopathic Parkinson disease. Am. J. Neuroradiol. 39, 1460–1467. https://doi.org/10.3174/ajnr.A5702.

Reimão, S., Pita Lobo, P., Neutel, D., Guedes, L.C., Coelho, M., Rosa, M.M., Azevedo, P., Ferreira, J., Abreu, D., Gonçalves, N., Nunes, R.G., Campos, J., Ferreira, J.J., 2015. Substantia nigra neuromelanin-MR imaging differentiates essential tremor from Parkinson's disease. Mov. Disord. 30, 953–959. https://doi.org/10.1002/mds.26182.

Reimão, S., Ferreira, S., Nunes, R.G., Pita Lobo, P., Neutel, D., Abreu, D., Gonçalves, N., Campos, J., Ferreira, J.J., 2016. Magnetic resonance correlation of iron content with neuromelanin in the substantia nigra of early-stage Parkinson's disease. Eur. J. Neurol. 23, 368–374. https://doi.org/10.1111/ene.12838.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 9351, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28.

Sasaki, M., Shibata, E., Tohyama, K., Takahashi, J., Otsuka, K., Tsuchiya, K., Takahashi, S., Ehara, S., Terayama, Y., Sakai, A., 2006. Neuromelanin magnetic resonance imaging of locus ceruleus and substantia nigra in Parkinson's disease. Neuroreport 17, 1215–1218. https://doi.org/10.1097/01.wnr.0000227984.84927.a7.

Schwarz, S.T., Rittman, T., Gontu, V., Morgan, P.S., Bajaj, N., Auer, D.P., 2011. T1-Weighted MRI shows stage-dependent substantia nigra signal loss in Parkinson's disease. Mov. Disord. 26, 1633–1638. https://doi.org/10.1002/mds.23722.

Schwarz, S.T., Xing, Y., Tomar, P., Bajaj, N., Auer, D.P., 2017. In Vivo assessment of brainstem depigmentation in Parkinson disease: Potential as a severity marker for

multicenter studies. Radiology 283, 789–798. https://doi.org/10.1148/RADIOL.2016160662.

Shinde, S., Prasad, S., Saboo, Y., Kaushick, R., Saini, J., Pal, P.K., Ingalhalikar, M., 2019. Predictive markers for Parkinson's disease using deep neural nets on neuromelanin sensitive MRI. NeuroImage Clin. 22, 101748 https://doi.org/10.1016/j.nicl.2019.101748.

Sudre, C.H., Li, W., Vercauteren, T., Ourselin, S., Jorge Cardoso, M., 2017. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 10553 LNCS, 240–248. 10.1007/978-3-319-67558-9_28.

Sulzer, D., Cassidy, C., Horga, G., Kang, U.J., Fahn, S., Casella, L., Pezzoli, G., Langley, J., Hu, X.P., Zucca, F.A., Isaias, I.U., Zecca, L., 2018. Neuromelanin detection by magnetic resonance imaging (MRI) and its promise as a biomarker for Parkinson's disease. npj Park Dis. 4, 11. https://doi.org/10.1038/s41531-018-0047-3.

Takahashi, H., Watanabe, Y., Tanaka, H., Mihara, M., Mochizuki, H., Liu, T., Wang, Y., Tomiyama, N., 2018. Quantifying changes in nigrosomes using quantitative susceptibility mapping and neuromelanin imaging for the diagnosis of early-stage Parkinson's disease. Br. J. Radiol. 91, 1–7. https://doi.org/10.1259/bjr.20180037.

Takahashi, H., Watanabe, Y., Tanaka, H., Mochizuki, H., Kato, H., Hatazawa, J., Tomiyama, N., 2019. Quantifying the severity of Parkinson disease by use of dopaminergic neuroimaging. Am. J. Roentgenol. 213, 163–168. https://doi.org/10.2214/AJR.18.20655.

Taniguchi, D., Hatano, T., Kamagata, K., Okuzumi, A., Oji, Y., Mori, A., Hori, M., Aoki, S., Hattori, N., 2018. Neuromelanin imaging and midbrain volumetry in progressive supranuclear palsy and Parkinson's disease. Mov. Disord. 33, 1488–1492. https://doi.org/10.1002/mds.27365.

Tournier, J.D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens, D., Jeurissen, B., Yeh, C.H., Connelly, A., 2019. MRtrix3: a fast, flexible and open software framework for medical image processing and visualisation. Neuroimage. https://doi.org/10.1016/j.neuroimage.2019.116137.

Varrone, A., Dickson, J.C., Tossici-Bolt, L., Sera, T., Asenbaum, S., Booij, J., Kapucu, O.L., Kluge, A., Knudsen, G.M., Koulibaly, P.M., Nobili, F., Pagani, M., Sabri, O., Vander Borght, T., Van Laere, K., Tatsch, K., 2013. European multicentre database of healthy controls for [123I]FP- CIT SPECT (ENC-DAT): age-related effects, gender differences and evaluation of different methods of analysis. Eur. J. Nucl. Med. Mol. Imaging 40, 213–227. https://doi.org/10.1007/s00259-012-2276-8.

Wang, J., Li, Y., Huang, Z., Wan, W., Zhang, Y., Wang, C., Cheng, X., Ye, F., Liu, K., Fei, G., Zeng, M., Jin, L., 2018. Neuromelanin-sensitive magnetic resonance imaging features of the substantia nigra and locus coeruleus in de novo Parkinson's disease and its phenotypes. Eur. J. Neurol. 25, 949–955. https://doi.org/10.1111/ene.13628.

Wang, J., Huang, Z., Li, Y., Ye, F., Wang, C., Zhang, Y., Cheng, X., Fei, G., Liu, K., Zeng, M., Zhong, C., Jin, L., 2019. Neuromelanin-sensitive MRI of the substantia nigra: an imaging biomarker to differentiate essential tremor from tremor-dominant Parkinson's disease. Park. Relat. Disord. 58, 3–8. https://doi.org/10.1016/j.parkreldis.2018.07.007.

Xing, Y., Sapuan, A., Dineen, R.A., Auer, D.P., 2018. Life span pigmentation changes of the substantia nigra detected by neuromelanin-sensitive MRI. Mov. Disord. 33, 1792–1799. https://doi.org/10.1002/mds.27502.

Xing, Y., Sapuan, A.H., Martín-Bastida, A., Naidu, S., Tench, C., Evans, J., Sare, G., Schwarz, S.T., Al-bachari, S., Parkes, L.M., Kanavou, S., Raw, J., Silverdale, M., Bajaj, N., Pavese, N., Burn, D., Piccini, P., Grosset, D.G., Auer, D.P., 2022. Neuromelanin-MRI to quantify and track nigral depigmentation in Parkinson's Disease: a multicenter longitudinal study using template-based standardized analysis. Mov. Disord. 37, 1028–1039. https://doi.org/10.1002/mds.28934.

Zecca, L., Fariello, R., Riederer, P., Sulzer, D., Gatti, A., Tampellini, D., 2002. The absolute concentration of nigral neuromelanin, assayed by a new sensitive method, increases throughout the life and is dramatically decreased in Parkinson's disease. FEBS Lett. 510, 216–220. https://doi.org/10.1016/S0014-5793(01)03269-0.

Zecca, L., Swartz, H.M., 1993. Total and paramagnetic metals in human substantia nigra and its neuromelanin. J. Neural Transm. Park. Dis. Dement. Sect. 5, 203–213. https://doi.org/10.1007/BF02257675.

Zhang, W., Phillips, K., Wielgus, A.R., Liu, J., Albertini, A., Zucca, F.A., Faust, R., Qian, S. Y., Miller, D.S., Chignell, C.F., Wilson, B., Jackson-Lewis, V., Przedborski, S., Joset, D., Loike, J., Hong, J.S., Sulzer, D., Zecca, L., 2011. Neuromelanin activates microglia and induces degeneration of dopaminergic neurons: Implications for progression of parkinson's disease. Neurotox. Res. 19, 63–72. https://doi.org/10.1007/s12640-009-9140-z.

Zucca, F.A., Basso, E., Cupaioli, F.A., Ferrari, E., Sulzer, D., Casella, L., Zecca, L., 2014. Neuromelanin of the human substantia Nigra: an update. Neurotox. Res. 25, 13–23. https://doi.org/10.1007/s12640-013-9435-y.

Zucca, F.A., Vanna, R., Cupaioli, F.A., Bellei, C., De Palma, A., Di Silvestre, D., Mauri, P., Grassi, S., Prinetti, A., Casella, L., Sulzer, D., Zecca, L., 2018. Neuromelanin organelles are specialized autolysosomes that accumulate undegraded proteins and lipids in aging human brain and are likely involved in Parkinson's disease. npj Park Dis. 4, 17. https://doi.org/10.1038/s41531-018-0050-8.