## RESEARCH

# Convolutional neural network optimizes the application of diffusion kurtosis imaging in Parkinson's disease

Junyan Sun[1†], Ruike Chen[2†], Qiqi Tong[2,3], Jinghong Ma[4], Linlin Gao[1], Jiliang Fang[5], Dongling Zhang[1], Piu Chan[1,6,7,8], Hongjian He[2,9*] and Tao Wu[1,6,7,8*]

## Abstract

**Objectives:** The literature regarding the use of diffusion-tensor imaging-derived metrics in the evaluation of Parkinson's disease (PD) is controversial. This study attempted to assess the feasibility of a deep-learning-based method for detecting alterations in diffusion kurtosis measurements associated with PD.

**Methods:** A total of 68 patients with PD and 77 healthy controls were scanned using scanner-A (3 T Skyra) (DATASET-1). Meanwhile, an additional five healthy volunteers were scanned with both scanner-A and an additional scanner-B (3 T Prisma) (DATASET-2). Diffusion kurtosis imaging (DKI) of DATASET-2 had an extra *b* shell compared to DATASET-1. In addition, a 3D-convolutional neural network (CNN) was trained from DATASET-2 to harmonize the quality of scalar measures of scanner-A to a similar level as scanner-B. Whole-brain unpaired *t* test and Tract-Based Spatial Statistics (TBSS) were performed to validate the differences between the PD and control groups using the model-fitting method and CNN-based method, respectively. We further clarified the correlation between clinical assessments and DKI results.

**Results:** An increase in mean diffusivity (MD) was found in the left substantia nigra (SN) in the PD group. In the right SN, fractional anisotropy (FA) and mean kurtosis (MK) values were negatively correlated with Hoehn and Yahr (H&Y) scales. In the putamen (Put), FA values were positively correlated with the H&Y scales. It is worth noting that these findings were only observed with the deep learning method. There was neither a group difference nor a correlation with clinical assessments in the SN or striatum exceeding the significance level using the conventional model-fitting method.

**Conclusions:** The CNN-based method improves the robustness of DKI and can help to explore PD-associated imaging features.

**Keywords:** Parkinson's disease, Diffusion kurtosis imaging, Convolutional neural network, Mean kurtosis, Kurtosis fractional anisotropy, Mean diffusivity

*Correspondence: hhezju@zju.edu.cn; wutao69@163.com
†Junyan Sun and Ruike Chen contributed equally to this work
1 Department of Neurobiology, Neurology and Geriatrics, Xuanwu Hospital of Capital Medical University, National Clinical Research Center for Geriatric Disease, Beijing 100053, China
2 Center for Brain Imaging Science and Technology, College of Biomedical Engineering and Instrumental Science, Zhejiang University, Hangzhou 310027, Zhejiang, China
Full list of author information is available at the end of the article

## 1 Introduction

Parkinson's disease (PD) is a common neurodegenerative disease characterized by bradykinesia, resting tremor, rigidity, postural balance disturbance, and non-motor manifestations [1]. Beyond the deficiency of dopaminergic neurons and aggregation of Lewy bodies in the basal ganglia, pathological changes in PD are associated with axonal lesions and synaptic dysfunction, which

Sun *et al. Brain Inf.*    (2021) 8:18

Page 2 of 12

contribute to the impairment of white matter integrity [2]. Given the limitation of discerning the intrinsic details and pathological heterogeneity in brain tissues, it is challenging to identify PD-associated microstructural changes using conventional magnetic resonance imaging (MRI). Diffusion-weighted MRI techniques, such as diffusion-tensor imaging (DTI), can non-invasively probe the microstructural properties via the diffusion of water molecules in vivo [3–7]. It has been reported that DTI-derived metrics, such as fractional anisotropy (FA) and mean diffusivity (MD), showed significant differences in the substantia nigra (SN) and some white matter areas between PD patients and controls [8]. Furthermore, diffusion kurtosis imaging (DKI), which is based on DTI and considers the non-Gaussian diffusion of water molecules, was reported as a more sensitive technique to evaluate the pathological characteristics of PD patients [3, 8].

However, previous studies have yielded inconsistent or controversial findings. For example, while some studies reported decreased FA, increased MD, and/or increased mean kurtosis (MK) values in the SN [6, 9–11], other studies [6, 12] observed increased FA values in the SN in PD patients. It has been reported that the FA value of the SN was higher in PD patients than in healthy controls (HCs) [11]. Additionally, Kamagata et al. found decreased MK and FA values of white matter, such as the inferior fronto-occipital fasciculus (IFOF), anterior corona radiata (ACR), and superior longitudinal fasciculi (SLF) [7], while Wen et al. showed increased FA values of the IFOF and bilateral SLF in tremor-dominant PD patients [13].

It is speculated that the heterogeneity of PD patients being recruited and various acquisition protocols of diffusion MRI scanning may have contributed to these controversial findings [4, 5, 8, 14]. DKI can sensitively reflect microstructural complexity, particularly in isotropic tissues such as the gray matter [15]. However, because the gray matter microstructure lacks evident directionality, diffusion-weighted imaging (DWI) signals can be easily affected by noise and limited spatial resolution [16], thus leading to inaccurate findings of alterations in DKI scalar measures. Another limitation to the wide application of DKI is that high $b$-value diffusion signals, which are required for the accurate calculation of DKI scalar measures, are often difficult to obtain in clinical settings.

Recently, deep learning, an important branch of machine learning, has shown significant potential for improving the performance of neuroimaging findings [17–21]. As one of the representative algorithms of deep learning, convolutional neural network (CNN) adopts convolution and down-sampling to certain layers with less computation; adjusts the network weights through the back-propagation and stochastic gradient descent algorithm; recognizes the features or patterns of the raw imaging inputs automatically; and then achieves the classification, identification, and prediction of inputs [21–23].

Li et al. [22] recently proposed a three-dimensional hierarchical CNN (3D H-CNN) to improve the estimation of DKI scalar measures from limited diffusion-weighted (DW) images. Three-dimensional convolution kernels were introduced to automatically extract and learn the features of the DW-images. Only part scalar measures were of clinical interest instead of the full tensors, and the 3D H-CNN (hereafter called CNN) method makes it possible to complete fast and optimized DKI acquisition within 1 min. This method also considers cross-voxel information, which was confirmed to provide enhanced efficiency for estimating DKI scalar measures and improved robustness against noise.

Therefore, in the current study, we aimed to use this CNN-based method to improve the estimation of DKI scalar measures and to determine whether the improved measures can help to delineate PD-associated imaging features.

## 2 Materials and methods

### 2.1 Participants

Sixty-eight patients with PD who met the Movement Disorder Society clinical diagnostic criteria for PD were recruited from the Movement Disorders Clinic of the Xuanwu Hospital of Capital Medical University. We recruited 77 HCs who met the following criteria: (1) aged ≥ 40 years; (2) no history of neurological or psychiatric diseases; (3) no family history of neurodegenerative disorders, and (4) no apparent cerebral lesions on structural MRI. The Movement Disorder Society Unified Parkinson's Disease Rating Scale, part III (MDS- UPDRS III) and Hoehn and Yahr (H&Y) scale were performed in all PD patients while they were in the off-state. Their demographic details are summarized in Table 1. In addition, seven healthy volunteers (M/F = 1/6, age = 26.4 ± 1.6 years) were recruited and their DKI

**Table 1** Demographic and clinical assessments of subjects

|  | PD ($n = 68$) | HC ($n = 77$) | $p$ value |
|---|---|---|---|
| Age in years, mean (SD) | 58.94 (8.969) | 59.58 (8.537) | 0.659 |
| Gender(M/F)[§] | 36/32 | 30/47 | 0.092 |
| Education in years, mean(SD) | 11.81 (3.316) | 11.58 (4.143) | 0.074 |
| H&Y score, median(range) | 2 (1–3) | 0 | NA |
| UPDRS III, mean (SD) | 26.5 (12.261) | NA | NA |
| Duration in years, Median (range) | 4 (0.5–20) | NA | NA |

*M* male, *F* female, *SD* standard variation, *NA* not applicable

[§] Pearson's Chi-square test

Sun *et al. Brain Inf.* (2021) 8:18

Page 3 of 12

data were collected using two different MRI scanners. This experiment was guided by and adhered to the Declaration of Helsinki and was approved by the Institutional Review Board of Xuanwu Hospital. All included participants have signed informed consent before the experiment.

## 2.2 MRI

### 2.2.1 DATASET-1

For all PD patients and HCs, MRI data were acquired using a 3-T scanner-A (MAGNETOM Skyra, Siemens, Germany) equipped with a 20-channel receiver head and neck joint coil (opening 16 channels). DW-images were obtained in axial orientation using single-shot spin-echo echo-planar imaging sequences (SE-EPI). Diffusion weightings of $b = 1000$ and 2000s/mm$^2$ were applied along 30 noncollinear sensitive gradient directions. One $b = 0$ image was acquired, resulting in a total of 61 DW-images. The other imaging parameters were as follows: repetition time (TR) = 5000 ms, echo time (TE) = 105 ms, resolution = $2 \times 2 \times 2$ mm, field of view (FOV) = $220 \times 220$ mm, number of slices = 68.

### 2.2.2 DATASET-2

Seven healthy volunteers were scanned using two different MRI scanners. We randomly included five volunteers in the CNN training dataset and assigned the remaining two volunteers to the test dataset. Scan 1 also used scanner-A and the same scanning process as used for DATASET-1. For implementation of the CNN, DW-images were also acquired in Scan 2 using a 3-T scanner-B (MAGNETOM Prisma, Siemens Germany) equipped with a 64-channel RF coil. DW-images were obtained using a simultaneous multi-slice diffusion echo-planar imaging sequence (SMS-EPI). Diffusion weightings of $b = 1000$, 2000, and 3000 s/mm$^2$ were applied in 30 gradient directions. Six $b = 0$ images and one $b = 0$ image with an opposite phase encoding direction were acquired, resulting in a total of 96 DW-images. The $b = 0$ image with the reversed phase encoding direction was used to correct the field inhomogeneity-induced distortion. The other imaging parameters were as follows: TR = 3000 ms, TE = 75 ms, resolution = 2 mm × 2 mm × 2 mm, FOV = 220 mm × 220 mm, number of slices = 68.

## 2.3 Image processing

### 2.3.1 Preprocessing

The preprocessing pipeline for both datasets was mainly based on FSL (FMRIB Software Library, University of Oxford, UK) [24]. *Rician* noise was removed using *dwidenoise* included in MRtrix 3 [25], followed by Gibbs-ring removal. Bulk head movement was corrected using FSL, which linearly aligned each diffusion-weighted image to the first $b = 0$ image. For DATASET-2 and Scan 2, distortion correction was also performed using the *topup* and *eddy* tools [26]. The susceptibility-induced off-resonance field map was first estimated by *topup* using a pair of non-DW ($b = 0$) images acquired with reversed phase encoding directions anterior–posterior and posterior–anterior (AP and PA). It was then fed into *eddy* to correct for eddy current and motion-induced distortion.
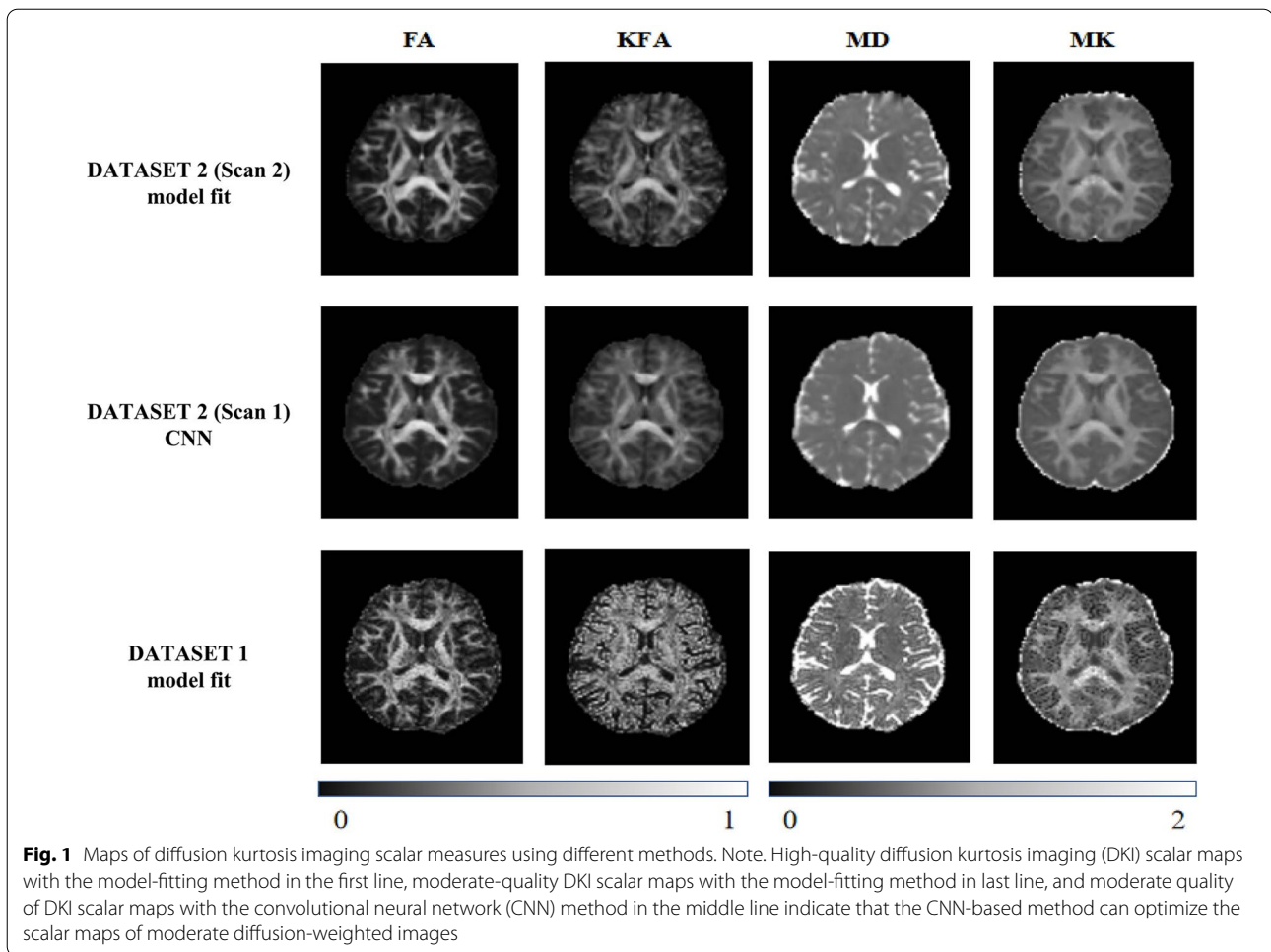
### 2.3.2 Model-fitting method

For DATASET-1 and Scan 2 of DATASET-2, the model-fitting method was conducted using DESIGNER (diffusion parameter EStImation with Gibbs and NoisE Removal, New York University, US), a post-processing pipeline capable of identifying and correcting various specific artifacts and confounding factors for improved accuracy, precision, and robustness compared to conventional linear least square method fitting (Fig. 1) [27].

### 2.3.3 CNN-based method

A CNN-based method was adopted to improve the estimation of DKI scalar measures from limited-quality DATASET-1 DW-images. The adopted CNN-based method included one input layer, several hidden layers, and two output layers. A dropout layer was inserted before each output layer to prevent overfitting. A $3 \times 3 \times 3$ convolution kernel was introduced in the first hidden layer to extract features from the preprocessed DW-images. The network was constructed using a hierarchical structure. The resulting DKI scalar measures were output through two different layers. The shallow output layer was connected to the penultimate hidden layer and was responsible for scalar measures (FA and MD). Kurtosis-related measures (MK and kurtosis FA) values were output through a deeper layer connected to the final hidden layer [22].

The pipeline of the CNN is shown in Fig. 2. Preprocessed DW-images of DATASET-2, Scan 1 were the training inputs of the CNN and the corresponding model-fitted MK, KFA, FA, and MD metrics for each healthy volunteer in DATASET-2, Scan 2 were defined as the ground truth. Scan 2 was registered to Scan 1 using FNIRT and FSL [28] for every volunteer before training. Specifically, we first removed non-brain tissue from images using BET2 and FSL. Nonlinear registration was then performed using FNIRT and FSL between the $b = 0$ images of Scans 1 and 2 for every volunteer. This resulted in registered images and corresponding nonlinear transformations. We then applied these transformations to the model-fitted DKI of Scan 2 to register them in the same space as Scan 1 images. We checked the registration results visually based on the alignment between the

Sun *et al. Brain Inf.* (2021) 8:18

Page 4 of 12



**Fig. 1** Maps of diffusion kurtosis imaging scalar measures using different methods. Note. High-quality diffusion kurtosis imaging (DKI) scalar maps with the model-fitting method in the first line, moderate-quality DKI scalar maps with the model-fitting method in last line, and moderate quality of DKI scalar maps with the convolutional neural network (CNN) method in the middle line indicate that the CNN-based method can optimize the scalar maps of moderate diffusion-weighted images

boundaries of the brain structures. We used the DKI data of five healthy volunteers to train the network. The training process was performed for 100 epochs.

We tested the trained CNN on two volunteers (not from the training dataset) and calculated the root-mean-squared error (RMSE) between the CNN-estimated or model-fitted DKI measures and the reference standard:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} \left(\hat{s}_i - s_i\right)^2}{N}}, \tag{1}$$

where $N$ is the number of voxels in the brain images, $\hat{s}_i$ is the CNN-estimated or model-fitted DKI scalar measure in the i$^{\text{th}}$ voxel, and $s_i$ is the ground truth DKI scalar measure in this voxel. Peak signal-to-noise ratios (PSNRs) were calculated as follows:

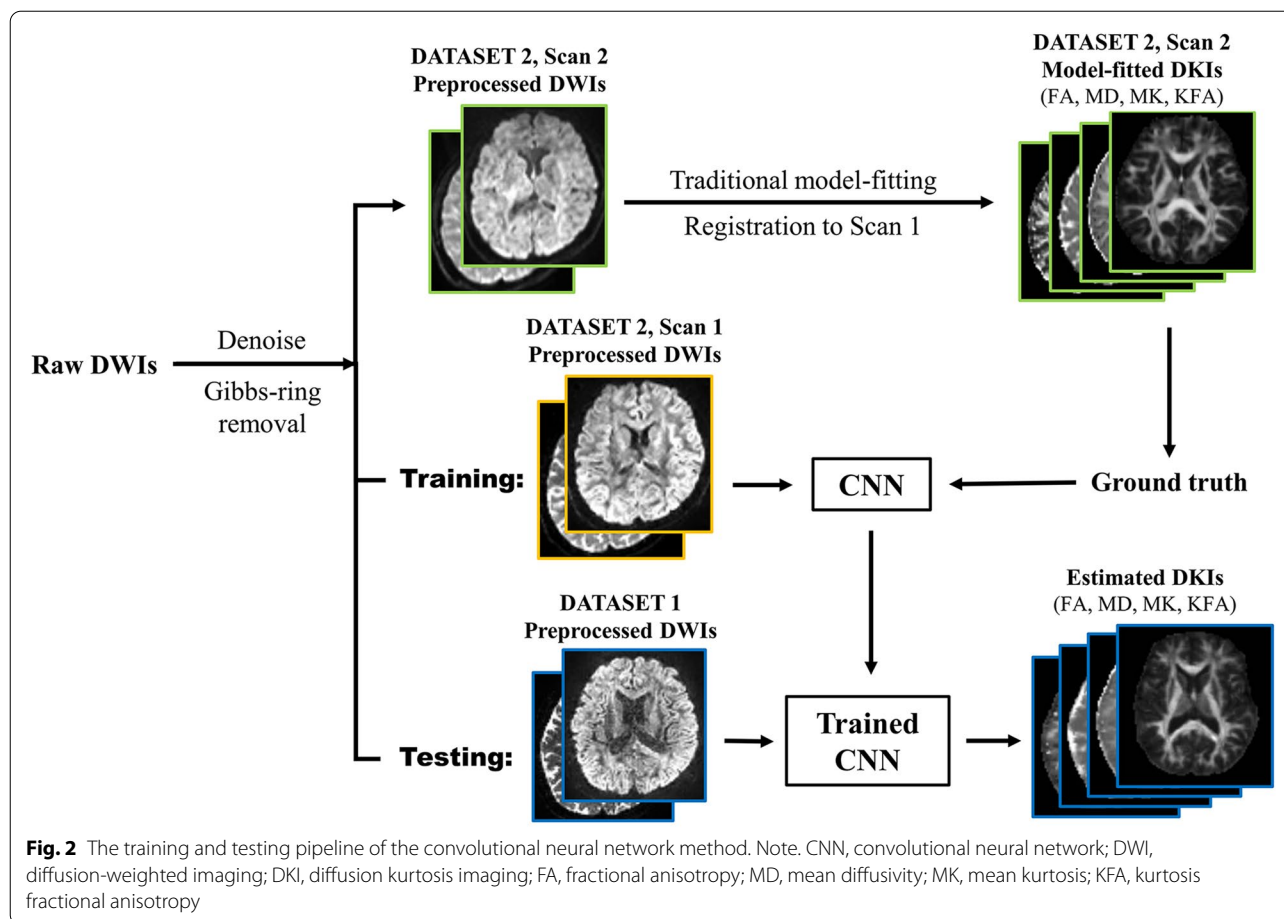$$\text{PSNR} = 10 \times \log_{10}\left(\frac{s_{MAX}^2}{MSE}\right), \tag{2}$$

where $s_{MAX}$ is the maximum signal value in a DKI scalar image and MSE is the mean-squared error, defined as:

$$\text{MSE} = \frac{\sum_{i=1}^{N} \left(\hat{s}_i - s_i\right)^2}{N}, \tag{3}$$

As shown in Fig. 3, the RMSE results of both testing subjects derived from the CNN-based method were lower than those of the conventional model-fitting method. The DKI scalar images estimated by the trained CNN showed a higher PSNR than the model-fitting method. With the same amount of diffusion-weighted signals acquired, the trained CNN was able to provide DKI estimations with higher qualities and closer to the ground truth.

We performed the same calculations on conventional model-fitted DKI results and compared them with the results of the CNN-based method. Other tests and validations on the robustness of the adopted 3D CNN structure have been described in a previous study [22].

Finally, we applied the trained CNN to subjects in both the HC and PD groups to estimate DKI measures, including FA, MD, KFA, and MK. The preprocessed

Sun *et al. Brain Inf.* (2021) 8:18

Page 5 of 12



**Fig. 2** The training and testing pipeline of the convolutional neural network method. Note. CNN, convolutional neural network; DWI, diffusion-weighted imaging; DKI, diffusion kurtosis imaging; FA, fractional anisotropy; MD, mean diffusivity; MK, mean kurtosis; KFA, kurtosis fractional anisotropy

DW-images from DATASET-1 as described in "Preprocessing" section were used as inputs.

## 2.4 Statistical analysis

The one-sample Shapiro–Wilk test was used to confirm the normality of each group's data. The Student's *t test* and Pearson Chi-square test were used to analyze age and sex variables, respectively.

### 2.4.1 Whole-brain unpaired t test

For the DW-images of PD patients and HCs, two datasets of DKI scalar measures were derived using model-fitting and CNN-based methods, respectively. FA maps were first registered to the Montreal Neurological Institute 152 (MNI152) standard space with a resolution of 1 mm isotropic using a combination of linear and nonlinear transforms. The resulting transformation was then applied to all other DKI maps for co-registration. For DKI scalar measures of both methods, whole-brain unpaired t-tests were performed to evaluate the ability of DKI to reveal differences between PD patients and HCs.

### 2.4.2 Tract-based spatial statistics

Furthermore, we performed Tract-Based Spatial Statistics (TBSS) analysis using the model-fitting and CNN-based method, respectively. TBSS extracted the mean FA maps to generate a white matter skeleton, realized by a tool for nonparametric permutation inference implemented in FSL [29].

The threshold-free cluster enhancement [30] based test was included in both the whole-brain unpaired t-test and TBSS analysis to improve robustness compared with conventional voxel-based tests. The number of permutation tests was set to 500 for both TBSS and t-test analyses. All significance thresholds were set at $p < 0.05$ and by family-wise error (FWE)-corrected.

To determine the specific brain regions to which the clusters with significance belonged, we utilized the FSL tool ATLASQUERY. This tool automatically matches the clusters to structural areas in user-specified atlas spaces and outputs the labels of the brain regions. In this study, we referred to parcellations from the Harvard–Oxford Cortical and Subcortical Structural Atlases [31] and JHU DTI-based white matter Atlases [32].

Sun *et al. Brain Inf.* (2021) 8:18

Page 6 of 12



**Fig. 3** The testing about the accuracy of CNN-estimated measures. Note. PSNR = peak signal-to-noise ratios; RMSE = root-mean-squared error
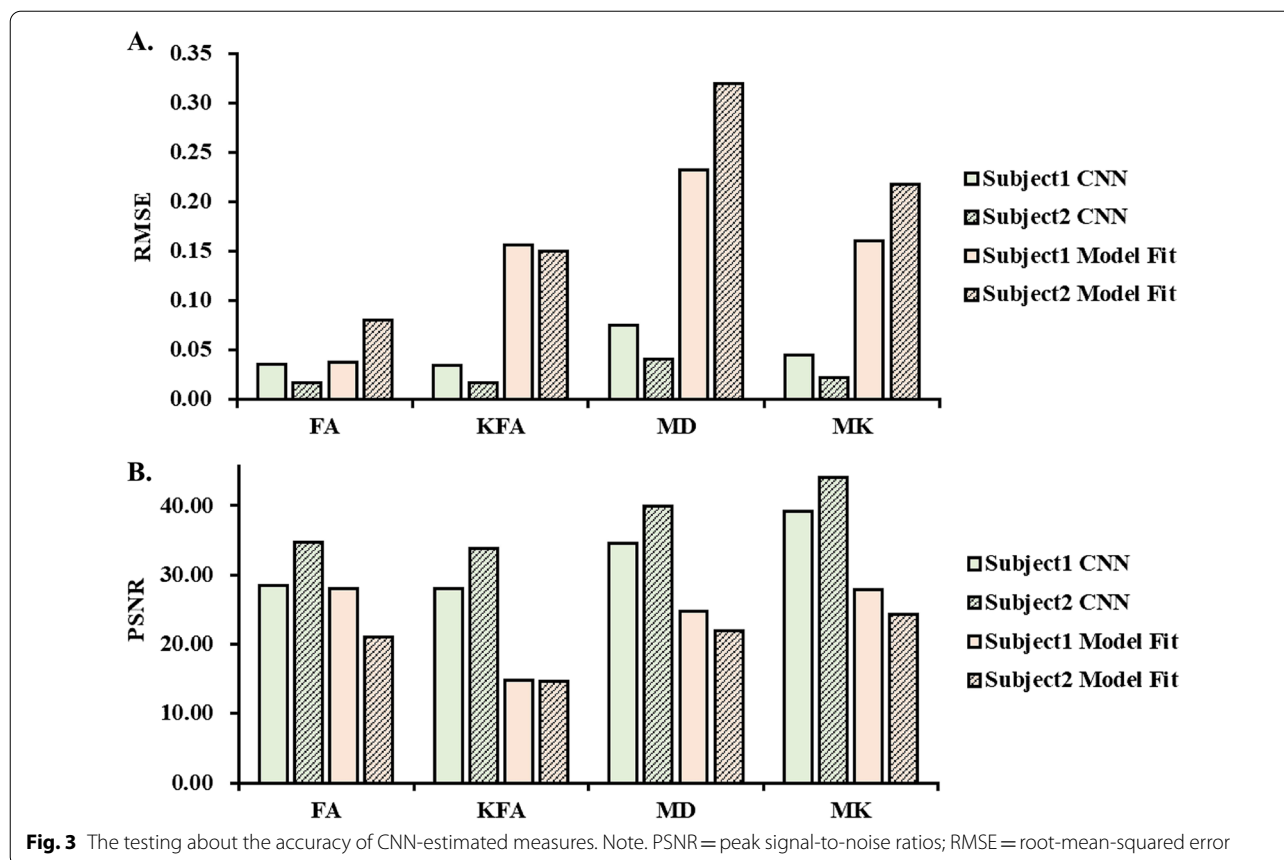
Notably, we determined whether there were intra-group differences in HCs by applying the whole-brain unpaired t-test and TBSS with model-fitted and CNN-based methods, respectively, before comparing differences between the groups.

### 2.4.3 Correlation analysis

To determine the clinical significance of DKI scalar measures using the CNN-based method more clearly, DKI scalar measures (MK, KFA, FA, and MD) determined using the model-fitting or CNN-based method showing significant between-group differences in the basal ganglia (SN, putamen, and caudate) were extracted and correlated with clinical assessments. Pearson's correlation analysis was used for normally distributed data, and Spearman's correlation analysis was used for non-normally distributed data. Correlations with significance were defined as $p < 0.016$ (Bonferroni-corrected). Statistical analyses were computed using IBM SPSS Statistics (version 25; IBM Corp., Armonk, NY, USA) and GraphPad Prism 8.0.1.

## 3 Results

### 3.1 CNN evaluations

As shown in Fig. 1, DKI scalar maps estimated by the CNN for DATASETS-1 had higher signal-to-noise ratios than those obtained by the model-fitting method. Higher-order DKI scalar measures such as KFA showed a clearer contrast between gray and white matter in the CNN-based results. For DATASET-2, the CNN-based results displayed almost the same quality as the reference standard (ground truth). The RMSE and PSNR results are listed in Fig. 3. The RMSE results of both testing subjects derived from the CNN-based method were lower than those from the conventional model-fitting method. The DKI scalar images estimated by the trained CNN showed a higher PSNR than the model-fitting method. With the same amount of diffusion-weighted signals acquired, the trained CNN was able to provide DKI estimations with higher qualities and closer to the ground truth.

### 3.2 Demographic features

Age ($p = 0.659$, Student's *t-test*), sex distribution ($p = 0.092$, Pearson Chi-square test), and education

Sun *et al. Brain Inf.*     (2021) 8:18

Page 7 of 12

($p = 0.074$, Student's *t-test*) did not differ between the PD and HC groups. The demographic details are summarized in Table 1.

### 3.3 Validation of intra-group differences between the PD and HC groups

Before comparing the PD patients and healthy controls, we separated all HCs and PD patients into two age-matched groups, respectively, and then conducted within-group comparisons, and found no significant difference with either the CNN-based or model-fitting method. It is preliminarily ruled out the possibility of false positives in the CNN-based method.

### 3.4 Whole-brain unpaired t-test analysis

#### 3.4.1 Model-fitting method

FA values in the bilateral putamen (Put) and globus pallidus (GP), left caudate (Cau) and accumbens (Acc), bilateral superior corona radiata (SCR) and anterior thalamic

radiation (ATR), et al. were higher in PD patients than in HCs.

MD values in bilateral Put, GP, Cau, thalamus (Thal), bilateral cerebral cortex and white matter, bilateral posterior thalamic radiation (PTR) and inferior longitudinal fasciculus (ILF) and genu of corpus callosum (GCC), et al. were significantly increased in PD patients compared to HCs.

There was no significant difference in the MK and KFA values between the groups ($p < 0.05$, FWE-corrected).

Please check Fig. 4 and Additional file 1: Table S1 for more details.

#### 3.4.2 CNN-based method

FA values were increased in left Put, bilateral cerebral cortex and white matter, bilateral SCR and left ACR, et al., while KFA values were increased only in the left ACR and cerebral cortex and white matter in PD patients compared to HCs.
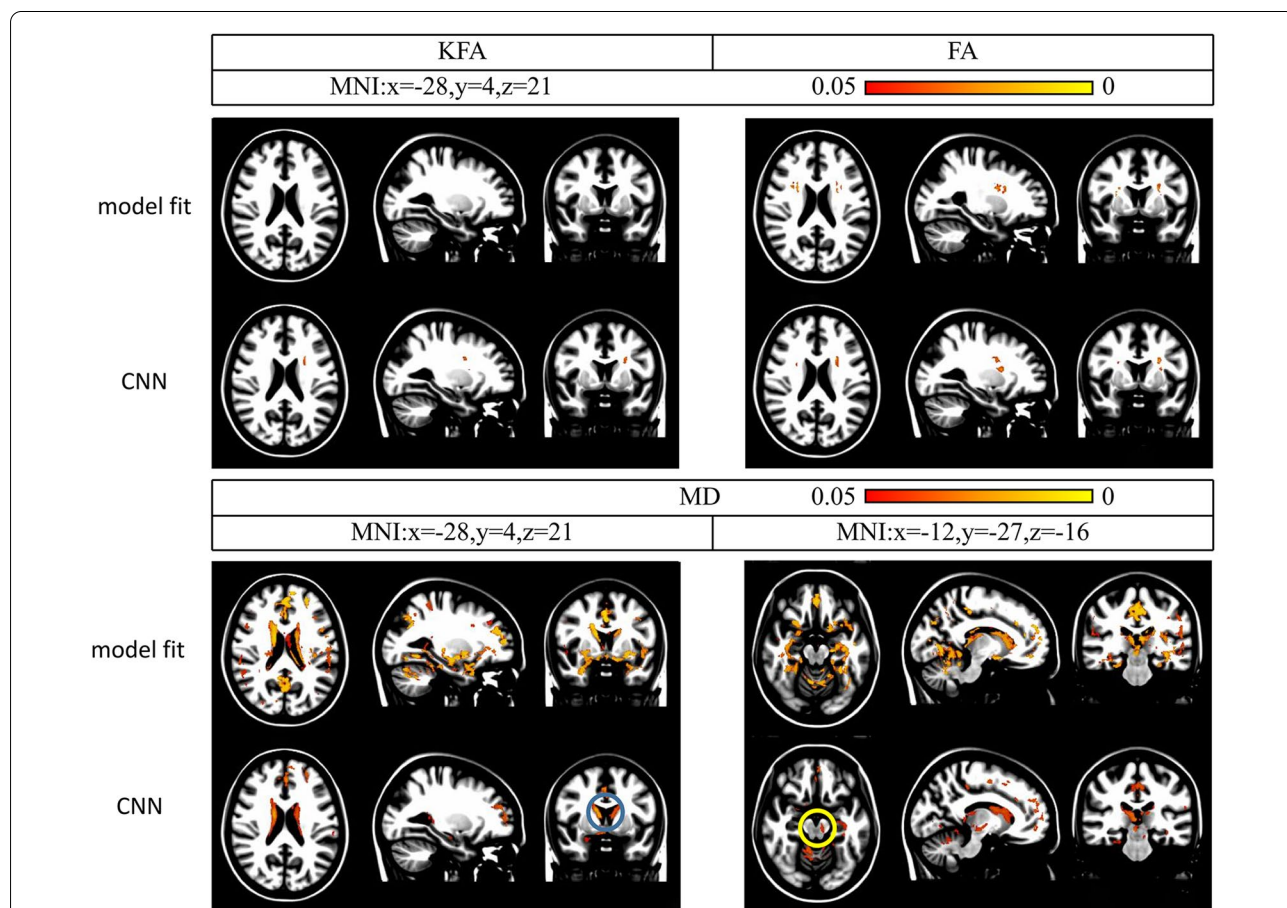


**Fig. 4** Whole-brain unpaired t-test analysis of diffusion kurtosis imaging measures using the CNN-based and model-fitting methods. Note. Increased mean diffusivity (MD) values in the bilateral caudate (Cau) (blue circle) and right substantia nigra (SN) (yellow circle) with the CNN-based method between the healthy control (HC) and Parkinson's disease (PD) groups ($p < 0.05$, family-wise error-corrected). KFA, kurtosis fractional anisotropy; FA, fractional anisotropy; MNI, Montreal Neurological Institute

Sun *et al. Brain Inf.* (2021) 8:18

Page 8 of 12

MD values in the left substantia nigra (SN) and left hippocampus and bilateral Cau and Thal, bilateral cerebral cortex and white matter, GCC and bilateral ATR, et al., were clearly increased in PD patients compared to HCs.

There was no significant difference in the MK values between the groups ($p < 0.05$, FWE-corrected).

Please check Fig. 4 and Additional file 1: Table S1 for more details.

### 3.5 TBSS analysis

#### 3.5.1 Model-fitting method

MK values were higher in the left ATR, left IFOF, left ILF, and left uncinate fasciculus (UNC), et al. in PD patients than in HCs.

FA values were increased in the left ATR, left IFOF, left corticospinal tract and left SLF, et al. in PD patients compared to HCs.

Compared to HCs, PD patients showed increased MD values in the bilateral IFOF, left ILF and bilateral UNC, et al.

There was no significant difference in the KFA values between the groups ($p < 0.05$, FWE-corrected).

Please check Fig. 5 and Additional file 2: Table S2 for more details.

#### 3.5.2 CNN-based method

MK values in the forceps minor, left IFOF, left UNC and left ATR, et al. were significantly increased in PD patients compared to HCs.

Further, PD patients showed higher KFA and FA values in multiple brain regions, such as the bilateral ATR, bilateral IFOF, bilateral SLF, and forceps minor et al. than HCs.

We did not find a significant difference in MD values between the two groups ($p < 0.05$, FWE-corrected).

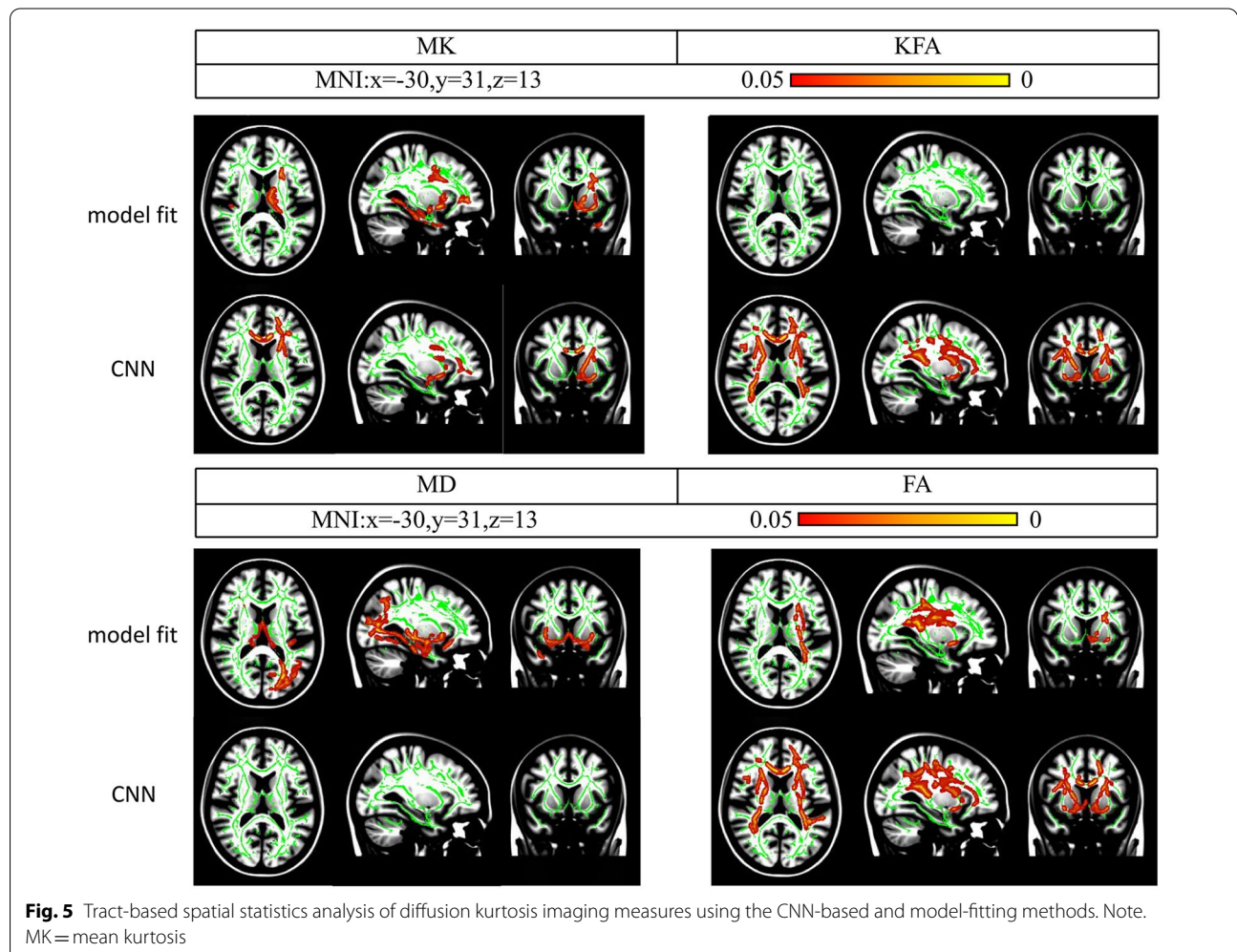Please check Fig. 5 and Additional file 2: Table S2 for more details.



**Fig. 5** Tract-based spatial statistics analysis of diffusion kurtosis imaging measures using the CNN-based and model-fitting methods. Note. MK = mean kurtosis

Sun *et al. Brain Inf.*   (2021) 8:18

Page 9 of 12

### 3.6 Correlation analysis

#### 3.6.1 Model-fitting method

We did not find any significant correlation between the DKI scalar measures and clinical assessments in patients with PD.

#### 3.6.2 CNN-based method

We found a positive correlation between the FA values of the left putamen and H&Y scales (r = 0.389, p = 0.001). A negative correlation was observed between the H&Y scales and FA and MK values in the right SN (r = − 0.390, p = 0.001; and r = − 0.349, p = 0.004, respectively) (Fig. 6).

### 4 Discussion

The major finding of this study was that the CNN-estimated MD values in the left SN and bilateral Cau were increased in PD patients compared to HCs. Additionally, the CNN-estimated FA and MK values in the right SN were negatively correlated with the H&Y scales, and CNN-estimated FA values in the left Put were positively correlated with the H&Y scales. In contrast, with the model-fitting method, there was no significant difference in MD values in the SN between PD patients and HCs, and there was no significant correlation between DKI scalar measures and clinical assessments in PD patients. Our findings suggest that the CNN-based method has the potential to optimize the estimation of DKI scalar measures and improve the sensitivity to detect PD-related imaging features.

In this study, we trained the CNN with data from five healthy volunteers and tested the trained network on another two healthy subjects. The RMSE and PSNR results suggested that the CNN-based method provided more accurate DKI scalar measures than the conventional model-fitting method. However, we failed to obtain ground truth data on PD patients due to limitations in clinical settings and the inconvenience for PD patients to travel. The rationale for applying a trained network to patient data is that the relationships between the original DW-images and the corresponding DKI scalar measures were learned based on voxel-wise diffusion data. That is,
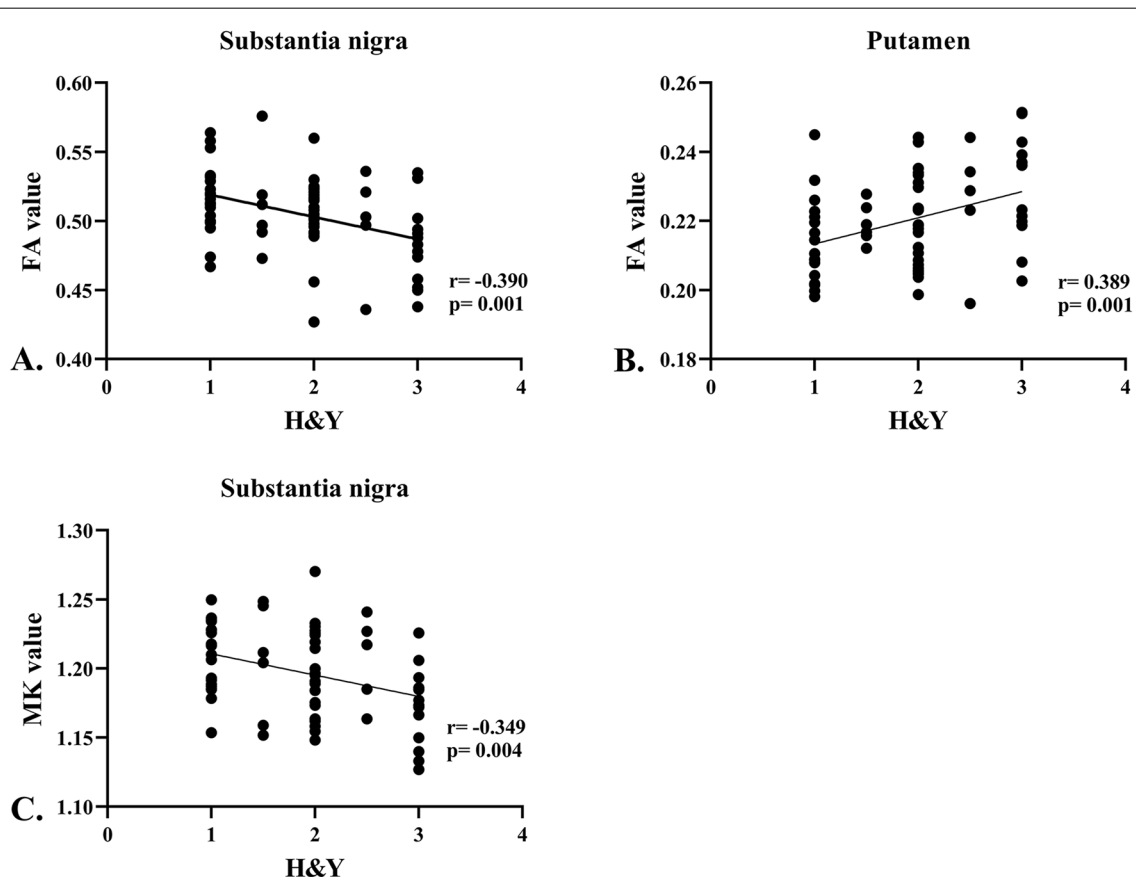


**Fig. 6** Spearman's correlation between the diffusion kurtosis imaging scalar measures and Hoehn and Yahr scales. **A**. Negative correlation between the fractional anisotropy (FA) values in the substantia nigra and Hoehn and Yahr (H&Y) scales. **B**. Positive correlation between the FA values in the putamen and H&Y scales. **C**. Negative correlation between the mean kurtosis values in the substantia nigra and H&Y scales

Sun *et al. Brain Inf.* (2021) 8:18

Page 10 of 12

the CNN provided estimations of DKI in every voxel and was sensitive to different diffusion signals. Therefore, it is natural to assume that the trained CNN was capable of estimating DKI scalar measures from another diffusion dataset that shared common acquisition parameters with the training dataset.

Using the CNN-based method, we found greater MD values in several brain regions in PD patients than in HCs, particularly in the left SN, which is consistent with previous reports using region of interest (ROI) analysis [10, 11, 14, 33–35]. PD is characterized by the progressive death of dopaminergic neurons in the SN, followed by the loss of dopaminergic projections from the SN to the striatum, resulting in a series of motor and non-motor symptoms [36]. According to the mathematical concept of a tensor, the three-dimensional shape of the diffusion elliptical structure depends on three eigenvalues ($\lambda 1$, $\lambda 2$, $\lambda 3$) of orthogonal principal axes without directions. The MD value was the average of the three eigenvalues. The impaired axons and neurons and loss of myelin integrity in PD patients result in a decrease in the restriction of water molecule displacement, which induces increased MD values [35, 37]. Regionally increased MD values in the left SN and bilateral Cau estimated by the CNN method were consistent with the pathological lesions in PD patients. In contrast, we did not find increased MD values in the SN in patients with PD compared to HCs by applying the model-fitting method. This finding indicates that the CNN-based method can better reveal the pathological features of PD than the model-fitting method.

We did not observe modulation of FA and MK values in the SN in PD patients, which is in line with a previous report [14]. In contrast, some previous studies based on ROI analysis showed decreased or increased FA and/or increased MK in the SN in PD patients [6, 34]. We speculate that different analysis methods may be responsible for these controversial results. The whole-brain unpaired t-test, moving beyond the hypothesis-driven ROI analysis, focused the statistical information on each voxel accompanied with increased partial volume effects and false-positive risk, particularly within the pathological brain tissues. Moreover, we suggest that these controversial findings may be due to the heterogeneity of recruited patients and variations in imaging quality [4, 8]. In addition, it has been reported that iron deposition could increase FA values and decrease MD values in white and gray matter [38]. Numerous reports have demonstrated iron accumulation in the SN [39–41]. Thus, different levels of iron deposition in the SN may also have contributed to these inconsistent findings.

We found a negative correlation between the H&Y scales and CNN-estimated FA values and MK values in the SN, as well as a positive correlation between the H&Y

scales and CNN-estimated FA values in the Put. These results indicated that FA and MK in the SN decreased, while FA in the Put increased with disease progression. As most of our patients were in the early stages (55 of our patients were at H&Y stages 1 and 2), it is possible to detect decreased FA in the SN if more advanced patients were enrolled. We did not find any significant correlation between DKI scalar measures and clinical assessments in PD patients using the model-fitting method, which further proves that using the CNN-based method to estimate DKI measures can improve the ability to explore PD-related neural modulations compared to using the model-fitting method.

For the TBSS analysis, increased FA values were observed in the brain white matter, such as ATR, IFOF with both methods, which was in line with previous studies [13, 42–44]. It has been shown that increased FA in these white matter regions correlates with better olfactory function and lower motor severity [45]. Thus, the increased diffusional properties of white matter might reflect microstructural compensation [45].

We observed greater MK values in the white matter in PD patients, which is inconsistent with previous reports. Previous studies found no significant difference in MK values [46, 47], or decreased MK values in the anterior cingulum, IFOF, and UNC in PD patients. We suggest that the heterogeneity of recruited patients and differences in the protocol of DW-image acquisition and image processing may have contributed to these inconsistent findings. In addition, we found increased KFA in the white matter, which has not been reported previously. KFA values, resembling the FA definition, quantify the degree of anisotropy of non-Gaussian diffusion. In the current study, the increased KFA and FA values were observed in the same white matter fibers. To date, only a small number of studies have focused on kurtosis changes in the white matter in PD patients [7, 48], and it is necessary to perform large cohort studies to elucidate the microstructural changes in white matter in PD patients.

There are some limitations to this study, such as the use of healthy volunteer subjects, which only provided DKI estimations closer to the ground truth. While we have validated the assumption that the CNN trained network is applicable to the PD datasets, further efforts will be made to include more ground truth data and more subjects to obtain more accurate results.

In conclusion, the CNN-based method has the potential to sensitively detect nigral pathology and improve the robustness and performance of DKI with few DW-images, and then to differentiate PD patents from HCs. In addition, compared with the model-fitting method, the CNN-based method can better determine the

Sun *et al. Brain Inf.*    (2021) 8:18

Page 11 of 12

relationship between DKI parameter measures and clinical assessment susceptibility. These findings confirm that the CNN can contribute to the determination of PD-associated imaging features.

## Abbreviations
Acc: Accumbens; ACR: Anterior corona radiata; AP: Anterior–posterior; ATR: Anterior thalamic radiation; Cau: Caudate; CNN: Convolutional neural network; DKI: Diffusion kurtosis imaging; DTI: Diffusion tensor imaging; DW-images: Diffusion-weighted images; DWI: Diffusion-weighted imaging; FA: Fractional anisotropy; FSL: FMRIB Software Library; FOV: Field of view; FWE: Family-wise error; GCC: Genu of corpus callosum; GP: Globus pallidus; H&Y: Hoehn and Yahr; HCs: Healthy controls; IFOF: Inferior fronto-occipital fasciculus; ILF: Inferior longitudinal fasciculus; KFA: Kurtosis fractional anisotropy; MD: Mean diffusivity; MDS-UPDRS III: Movement Disorder Society Unified Parkinson's Disease Rating Scale, part III; MK: Mean kurtosis; MNI152: Montreal Neurological Institute 152; MRI: Magnetic resonance imaging; PA: Posterior–anterior; PD: Parkinson's disease; PSNR: Peak signal-to-noise ratio; Put: Putamen; RF: Radiofrequency; RMSE: Root-mean-squared error; ROI: Region of interest; SCR: Superior corona radiata; SE-EPI: Spin-echo echo-planar imaging sequences; SLF: Superior longitudinal fasciculi; SMS-EPI: Simultaneous multi-slice diffusion echo-planar imaging sequence; SN: Substantia nigra; TBSS: Tract-based spatial statistics; TE: Echo time; TFCE: Threshold-free cluster enhancement; Thal: Thalamus; TR: Repetition time.

## Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s40708-021-00139-z.

> **Additional file 1: Table S1.** Summary of the t-test results of the comparison between the two groups.
>
> **Additional file 2: Table S2.** Summary of the TBSS results of the comparison between the two groups.

## Availability of data and materials
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Declarations

## Competing interests
The authors declare that they have no competing interests.

## Author details
[1]Department of Neurobiology, Neurology and Geriatrics, Xuanwu Hospital of Capital Medical University, National Clinical Research Center for Geriatric Disease, Beijing 100053, China. [2]Center for Brain Imaging Science and Technology, College of Biomedical Engineering and Instrumental Science, Zhejiang University, Hangzhou 310027, Zhejiang, China. [3]Research Center for Healthcare Data Science, Zhejiang Lab, Hangzhou, Zhejiang, China. [4]Department of Neurology, Xuanwu Hospital of Capital Medical University, Beijing, China. [5]Department of Radiology, Guang'anmen Hospital, China Academy of Chinese Medical Sciences, Beijing, China. [6]Clinical Center for Parkinson's Disease, Capital Medical University, Beijing, China. [7]Key Laboratory for Neurodegenerative Disease of the Ministry of Education, Beijing Key Laboratory for Parkinson's Disease, Parkinson Disease Center of Beijing Institute for Brain Disorders, Beijing, China. [8]National Clinical Research Center for Geriatric Disorders, Beijing, China. [9]Key Laboratory for Biomedical Engineering of Ministry of Education, Zhejiang University, Hangzhou 310027, Zhejiang, China.

## References
1. Poewe W, Seppi K, Tanner CM et al (2017) Parkinson disease. Nat Rev Dis Primers 3:17013. https://doi.org/10.1038/nrdp.2017.13
2. Wong YC, Luk K, Purtell K et al (2019) Neuronal vulnerability in Parkinson disease: should the focus be on axons and synaptic terminals? Mov Disord 34:1406–1422. https://doi.org/10.1002/mds.27823
3. Jensen JH, Helpern JA, Ramani A, Lu H, Kaczynski K (2005) Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. Magn Reson Med 53:1432–1440. https://doi.org/10.1002/mrm.20508
4. Atkinson-Clement C, Pinto S, Eusebio A, Coulon O (2017) Diffusion tensor imaging in Parkinson's disease: review and meta-analysis. Neuroimage Clin 16:98–110. https://doi.org/10.1016/j.nicl.2017.07.011
5. Cochrane CJ, Ebmeier KP (2013) Diffusion tensor imaging in parkinsonian syndromes: a systematic review and meta-analysis. Neurology 80:857–864. https://doi.org/10.1212/WNL.0b013e318284070c
6. Wang JJ, Lin WY, Lu CS et al (2011) Parkinson disease: diagnostic utility of diffusion kurtosis imaging. Radiology 261:210–217. https://doi.org/10.1148/radiol.11102277
7. Kamagata K, Tomiyama H, Hatano T et al (2014) A preliminary diffusional kurtosis imaging study of Parkinson disease: comparison with conventional diffusion tensor imaging. Neuroradiology 56:251–258. https://doi.org/10.1007/s00234-014-1327-1
8. Bergamino M, Keeling EG, Mishra VR, Stokes AM, Walsh RR (2020) Assessing white matter pathology in early-stage Parkinson disease using diffusion MRI: a systematic review. Front Neurol 11:314. https://doi.org/10.3389/fneur.2020.00314
9. Vaillancourt DE, Spraker MB, Prodoehl J et al (2009) High-resolution diffusion tensor imaging in the substantia nigra of de novo Parkinson disease. Neurology 72:1378–1384. https://doi.org/10.1212/01.wnl.0000340982.01727.6e
10. Joshi N, Rolheiser TM, Fisk JD et al (2017) Lateralized microstructural changes in early-stage Parkinson's disease in anterior olfactory structures, but not in substantia nigra. J Neurol 264:1497–1505. https://doi.org/10.1007/s00415-017-8555-3
11. Loane C, Politis M, Kefalopoulou Z et al (2016) Aberrant nigral diffusion in Parkinson's disease: a longitudinal diffusion tensor imaging study. Mov Disord 31:1020–1026. https://doi.org/10.1002/mds.26606
12. Lenfeldt N, Larsson A, Nyberg L, Birgander R, Forsgren L (2015) Fractional anisotropy in the substantia nigra in Parkinson's disease: a complex picture. Eur J Neurol 22:1408–1414. https://doi.org/10.1111/ene.12760
13. Wen MC, Heng HSE, Lu Z et al (2018) Differential white matter regional alterations in motor subtypes of early drug-naive Parkinson's disease patients. Neurorehabil Neural Repair 32:129–141. https://doi.org/10.1177/1545968317753075
14. Schwarz ST, Abaei M, Gontu V, Morgan PS, Bajaj N, Auer DP (2013) Diffusion tensor imaging of nigral degeneration in Parkinson's disease: a region-of-interest and voxel-based study at 3 T and systematic review

with meta-analysis. Neuroimage Clin 3:481–488. https://doi.org/10.1016/j.nicl.2013.10.006

15. Wang Q, Xu X, Zhang M (2010) Normal aging in the basal ganglia evaluated by eigenvalues of diffusion tensor imaging. AJNR Am J Neuroradiol 31:516–520. https://doi.org/10.3174/ajnr.A1862

16. Hansen B, Shemesh N, Jespersen SN (2016) Fast imaging of mean, axial and radial diffusion kurtosis. Neuroimage 142:381–393. https://doi.org/10.1016/j.neuroimage.2016.08.022

17. Liu M, Zhang J, Lian C, Shen D (2020) Weakly supervised deep learning for brain disease prognosis using MRI and incomplete clinical scores. IEEE Trans Cybern 50:3381–3392. https://doi.org/10.1109/tcyb.2019.2904186

18. Shamir RR, Duchin Y, Kim J et al (2019) Microelectrode recordings validate the clinical visualization of subthalamic-nucleus based on 7T magnetic resonance imaging and machine learning for deep brain stimulation surgery. Neurosurgery 84:749–757. https://doi.org/10.1093/neuros/nyy212

19. Zou L, Zheng J, Miao C, Mckeown MJ, Wang ZJ (2017) 3D CNN based automatic diagnosis of attention deficit hyperactivity disorder using functional and structural MRI. IEEE Access 5:23626–23636. https://doi.org/10.1109/ACCESS.2017.2762703

20. Yuan L, Wei X, Shen H, Zeng L-L, Hu D (2018) Multi-center brain imaging classification using a novel 3D CNN approach. IEEE Access. https://doi.org/10.1109/ACCESS.2018.2868813

21. Liu Z, Cao C, Ding S, Liu Z, Han T, Liu S (2018) Towards clinical diagnosis: automated stroke lesion segmentation on multi-spectral MR image using convolutional neural network. IEEE Access 6:57006–57016. https://doi.org/10.1109/ACCESS.2018.2872939

22. Li Z, Gong T, Lin Z et al (2019) Fast and robust diffusion kurtosis parametric mapping using a three-dimensional convolutional neural network. IEEE Access 7:71398–71411. https://doi.org/10.1109/access.2019.2919241

23. LeCun Y, Boser B, Denker JS et al (1989) Backpropagation applied to handwritten zip code recognition. Neural Comput 1:541–551. https://doi.org/10.1162/neco.1989.1.4.541

24. Woolrich MW, Jbabdi S, Patenaude B et al (2009) Bayesian analysis of neuroimaging data in FSL. Neuroimage 45:S173–S186. https://doi.org/10.1016/j.neuroimage.2008.10.055

25. Tournier JD, Smith R, Raffelt D et al (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

26. Andersson JLR, Sotiropoulos SN (2016) An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. Neuroimage 125:1063–1078. https://doi.org/10.1016/j.neuroimage.2015.10.019

27. Ades-Aron B, Veraart J, Kochunov P et al (2018) Evaluation of the accuracy and precision of the diffusion parameter EStimation with Gibbs and NoisE removal pipeline. Neuroimage 183:532–543. https://doi.org/10.1016/j.neuroimage.2018.07.066

28. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM (2012) FSL. Neuroimage 62:782–790. https://doi.org/10.1016/j.neuroimage.2011.09.015

29. Smith SM, Jenkinson M, Johansen-Berg H et al (2006) Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. Neuroimage 31:1487–1505. https://doi.org/10.1016/j.neuroimage.2006.02.024

30. Smith SM, Nichols TE (2009) Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. Neuroimage 44:83–98. https://doi.org/10.1016/j.neuroimage.2008.03.061

31. Desikan RS, Segonne F, Fischl B et al (2006) An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31:968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021

32. Hua K, Zhang J, Wakana S et al (2008) Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. Neuroimage 39:336–347. https://doi.org/10.1016/j.neuroimage.2007.07.053

33. Du G, Lewis MM, Sen S et al (2012) Imaging nigral pathology and clinical progression in Parkinson's disease. Mov Disord 27:1636–1643. https://doi.org/10.1002/mds.25182

34. Zhang G, Zhang Y, Zhang C et al (2015) Diffusion kurtosis imaging of substantia nigra is a sensitive method for early diagnosis and disease evaluation in Parkinson's disease. Parkinsons Dis 2015:207624. https://doi.org/10.1155/2015/207624

35. Duncan GW, Firbank MJ, Yarnall AJ et al (2016) Gray and white matter imaging: a biomarker for cognitive impairment in early Parkinson's disease? Mov Disord 31:103–110. https://doi.org/10.1002/mds.26312

36. Braak H, Rüb U, Gai WP, Del Tredici K (2003) Idiopathic Parkinson's disease: possible routes by which vulnerable neuronal types may be subject to neuroinvasion by an unknown pathogen. J Neural Transm 110:517–536. https://doi.org/10.1007/s00702-002-0808-2

37. Concha L (2014) A macroscopic view of microstructure: using diffusion-weighted images to infer damage, repair, and plasticity of white matter. Neuroscience 276:14–28. https://doi.org/10.1016/j.neuroscience.2013.09.004

38. Xu X, Wang Q, Zhong J, Zhang M (2015) Iron deposition influences the measurement of water diffusion tensor in the human brain: a combined analysis of diffusion and iron-induced phase changes. Neuroradiology 57:1169–1178. https://doi.org/10.1007/s00234-015-1579-4

39. Martin WR, Wieler M, Gee M (2008) Midbrain iron content in early Parkinson disease: a potential biomarker of disease status. Neurology 70:1411–1417. https://doi.org/10.1212/01.wnl.0000286384.31050.b5

40. Ward RJ, Zucca FA, Duyn JH, Crichton RR, Zecca L (2014) The role of iron in brain ageing and neurodegenerative disorders. Lancet Neurol 13:1045–1060. https://doi.org/10.1016/s1474-4422(14)70117-6

41. Straub S, Schneider TM, Emmerich J et al (2017) Suitable reference tissues for quantitative susceptibility mapping of the brain. Magn Reson Med 78:204–214. https://doi.org/10.1002/mrm.26369

42. Chen NK, Chou YH, Sundman M et al (2018) Alteration of diffusion-tensor magnetic resonance imaging measures in brain regions involved in early stages of Parkinson's disease. Brain Connect 8:343–349. https://doi.org/10.1089/brain.2017.0558

43. Taylor KI, Sambataro F, Boess F, Bertolino A, Dukart J (2018) Progressive decline in gray and white matter integrity in de novo Parkinson's disease: an analysis of longitudinal Parkinson progression markers initiative diffusion tensor imaging data. Front Aging Neurosci 10:318. https://doi.org/10.3389/fnagi.2018.00318

44. Wen MC, Heng HS, Ng SY, Tan LC, Chan LL, Tan EK (2016) White matter microstructural characteristics in newly diagnosed Parkinson's disease: an unbiased whole-brain study. Sci Rep 6:35601. https://doi.org/10.1038/srep35601

45. Sanjari Moghaddam H, Dolatshahi M, Mohebi F, Aarabi MH (2020) Structural white matter alterations as compensatory mechanisms in Parkinson's disease: a systematic review of diffusion tensor imaging studies. J Neurosci Res 98:1398–1416. https://doi.org/10.1002/jnr.24617

46. Surova Y, Nilsson M, Lampinen B et al (2018) Alteration of putaminal fractional anisotropy in Parkinson's disease: a longitudinal diffusion kurtosis imaging study. Neuroradiology 60:247–254. https://doi.org/10.1007/s00234-017-1971-3

47. Surova Y, Lampinen B, Nilsson M et al (2016) Alterations of diffusion kurtosis and neurite density measures in deep grey matter and white matter in Parkinson's disease. PLoS ONE 11:e0157755. https://doi.org/10.1371/journal.pone.0157755

48. Kamagata K, Tomiyama H, Motoi Y et al (2013) Diffusional kurtosis imaging of cingulate fibers in Parkinson disease: comparison with conventional diffusion tensor imaging. Magn Reson Imaging 31:1501–1506. https://doi.org/10.1016/j.mri.2013.06.009

## Publisher's Note