

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- |                                     |                                     |  |
|-------------------------------------|-------------------------------------|--|
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A description of all covariates tested   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | For null hypothesis testing, the test statistic (e.g. $F$ , $t$ , $r$ ) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br><i>Give <math>P</math> values as exact values whenever suitable.</i>                            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | Estimates of effect sizes (e.g. Cohen's $d$ , Pearson's $r$ ), indicating how they were calculated   |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

### Software and code

Policy information about [availability of computer code](#)

Data collection This study utilised three open-source datasets for model development and validation. Hence, no data was collected.

Data analysis All code uploaded to github and cited as fully open source project

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

This study utilised three open-source datasets for model development and validation: the Parkinson's Progression Markers Initiative (PPMI) MRI dataset, the NEUROCON dataset, and the Tao Wu dataset. All datasets are publicly available and can be accessed through their respective online repositories: <https://www.ppmi-info.org> and [https://fcon\\_1000.projects.nitrc.org/indi/retro/parkinsons.html](https://fcon_1000.projects.nitrc.org/indi/retro/parkinsons.html). All code for model implementation and analysis is available at <https://github.com/salilp42/KAN-MRI>.

## Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender	Sex and gender information was collected and reported in all three datasets. Total participants included 70 males and 72 females across all datasets (PPMI: 26 males, 33 females; NEUROCON: 21 males, 22 females; Tao Wu: 23 males, 17 females). This information was based on biological sex as reported in medical records. No specific sex/gender-based analyses were performed as this was not a primary aim of the study.
Population characteristics	Study participants included PD patients and healthy controls across three datasets. Mean ages were: PPMI 63.5 years (SD 11.1), NEUROCON 68.3 years (SD 11.0), and Tao Wu 65.0 years (SD 5.0). Disease duration varied: PPMI 1.7 years (SD 0.8), NEUROCON 4.8 years (SD 6.2), and Tao Wu 5.4 years (SD 3.9). PPMI included only newly diagnosed, untreated PD patients, while NEUROCON and Tao Wu included patients on dopaminergic medications.
Recruitment	All data was obtained from publicly available datasets: PPMI, NEUROCON, and Tao Wu. The PPMI dataset specifically recruited newly diagnosed PD patients within 2 years of diagnosis and without medication. This selection criterion may limit generalizability to broader PD populations.
Ethics oversight	As this study utilized existing public datasets, original ethical approvals were obtained by the primary data collectors (PPMI, NEUROCON, and Tao Wu research teams). Our analysis of these public datasets was conducted in accordance with their data usage agreements.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For 2D analyses, sample sizes were PPMI (n=5900 slices), Tao Wu (n=4000 slices), and NEUROCON (n=4300 slices), providing statistical power >0.99 to detect effect sizes of Cohen's d = 0.50. 3D analyses used smaller samples due to computational constraints (PPMI: n=59, Tao Wu: n=40, NEUROCON: n=43), achieving power of 0.69-0.85 for large effect sizes (Cohen's d = 0.80). Power calculations were performed prior to analysis to confirm adequacy of sample sizes.
Data exclusions	For PPMI dataset, only T1-weighted MRI scans that adhered to the PPMI imaging protocol were included. For 2D analysis, only axial slices centered on the midbrain (100 slices per subject) were selected. These criteria were established prior to analysis. No subjects were excluded from the NEUROCON or Tao Wu datasets.
Replication	Model performance was validated using both within-dataset cross-validation and cross-dataset hold-out testing. For 2D models, five-fold stratified group cross-validation was used. For 3D models, Leave-One-Out Cross-Validation was employed due to smaller sample sizes. All experiments were computationally deterministic and reproducible using the provided code.
Randomization	This study analyzed existing datasets where group allocation (PD vs Control) was predetermined by clinical diagnosis. As this was a retrospective analysis of pre-collected data, randomization was not applicable.
Blinding	All analyses were performed computationally using pre-labeled data from public repositories. The automated nature of the analysis pipeline meant investigator blinding was not relevant to model training or evaluation.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials &amp; experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

## Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input type="checkbox"/>	<input checked="" type="checkbox"/> MRI-based neuroimaging

## Magnetic resonance imaging

## Experimental design

Design type	Not applicable - this study analyzed existing structural MRI data from public repositories. No task-based or event-related imaging was performed.
Design specifications	Not applicable - this was a retrospective analysis of single timepoint structural MRI scans from established datasets.
Behavioral performance measures	Not applicable - no behavioral measures were collected as this study utilized existing structural MRI scans for diagnostic classification.

## Acquisition

Imaging type(s)	Structural T1-weighted MRI
Field strength	Variable across datasets. NEUROCON and Tao Wu used 3T; PPMI data collection used multiple scanners (field strengths documented in original PPMI protocols).
Sequence & imaging parameters	T1-weighted MPRAGE sequences were used across all datasets. TR/TE for NEUROCON: 1940ms/3.08ms; Tao Wu: 1100ms/3.39ms. PPMI parameters varied by collection site.
Area of acquisition	Whole brain T1-weighted structural scans
Diffusion MRI	<input type="checkbox"/> Used <input checked="" type="checkbox"/> Not used

## Preprocessing

Preprocessing software	Custom preprocessing pipeline including slice extraction (2D analysis) and volume resampling (3D analysis). 2D: 224x224 pixels; 3D: 128x128x128 voxels.
Normalization	Intensity normalization applied to scale values to [0,1] range for both 2D and 3D analyses.
Normalization template	Not applicable - only intensity normalization was performed; no spatial normalization was required for our analysis approach.
Noise and artifact removal	2D/3D Gaussian filtering ( $\sigma = 1\text{mm}$ ) applied for noise reduction.
Volume censoring	Not applicable - structural MRI analysis did not require volume censoring.

## Statistical modeling &amp; inference

Model type and settings	Deep learning models (ConvKAN, CNN, GCN) in both 2D and 3D implementations for binary classification (PD vs. control).
Effect(s) tested	Binary classification performance (accuracy, AUC, F1 score, recall) between PD and control groups.
Specify type of analysis:	<input checked="" type="checkbox"/> Whole brain <input type="checkbox"/> ROI-based <input type="checkbox"/> Both
Statistic type for inference (See <a href="#">Eklund et al. 2016</a> )	Not applicable - this study used deep learning models for whole-brain classification rather than traditional voxel-wise or cluster-wise inference methods.
Correction	Benjamini-Hochberg procedure was used to control false discovery rate in multiple comparisons of model performance metrics.

## Models & analysis

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Functional and/or effective connectivity
<input type="checkbox"/>	<input checked="" type="checkbox"/> Graph analysis
<input type="checkbox"/>	<input checked="" type="checkbox"/> Multivariate modeling or predictive analysis

### Graph analysis

Graph Convolutional Networks used k-nearest neighbors (k=6) for edge construction based on Euclidean distances between centroids. For 2D: 1000 superpixels per slice; for 3D: 1000 supervoxels per volume. Node features included intensity values and spatial information.

### Multivariate modeling and predictive analysis

Deep learning models (ConvKAN, CNN, GCN) performed binary classification (PD vs control) using whole-brain T1-weighted MRI data. Features were learned hierarchically through model architectures. Evaluation metrics included accuracy, AUC, F1 score, and recall, validated through cross-validation and hold-out testing.