

## 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](#).

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study. For final submission: please carefully check your responses for accuracy; you will not be able to make changes later.

### 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 checked="" type="checkbox"/> | <input 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

Gait videos were collected using the smartphone's built-in camera software with a minimum resolution of 720P and a minimum frame rate of 25 fps.

Data analysis

Data analysis was performed using Python along with the following packages: matplotlib (v3.7.5), moviepy (v1.0.3), numpy (v1.24.4), onnx (v1.17.0), onnxruntime-gpu (v1.19.0), opencord (v0.1.7), ptitprince (v0.2.7), scipy (v1.10.1), and seaborn (v0.11.0). The code for data preprocessing, analysis, model training, and evaluation is available at Code Ocean, which is associated with the article.

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](#)

We provided a toy dataset containing skeleton data, gait features extracted from the skeleton data and features obtained using the DMGrad-CAM method. This dataset is available at Code Ocean. To protect participants' privacy, the original gait videos are not provided. Detailed usage instructions are available in the capsule's README.md.

## Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	Sex was recorded during the data collection process and is reported in Table I of the paper.
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	A total of 87 patients with Parkinson's disease (PD) and 31 healthy control (HC) subjects were included in the study, comprising 55 females and 63 males (mean [SD] age 62.9 [7.4] years, mean [SD] height 166.2 [8.4] cm). Among the patients with PD, 19 participated in the medication response test, with a mean [SD] time since diagnosis of 5.4 [3.8] years.
Recruitment	Participants were recruited from Tianjin Huanhu Hospital in Tianjin, China. They were screened for Parkinson's disease, and their clinical tests were assessed by expert neurologists specializing in movement disorders. The inclusion criterion for the patient group was a confirmed diagnosis of Parkinson's disease and the capability to walk without assistance. Healthy controls, matched by age, were also included in the study.
Ethics oversight	All ethical and experimental procedures and protocols were approved by the Institutional Review Board of Tianjin Huanhu Hospital, Tianjin, China.

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://www.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	The sample size was determined by the availability and eligibility of individuals diagnosed with Parkinson's disease for participation in the study.
Data exclusions	Twelve participants were excluded from the study due to needing walking assistance (n=3), non-compliance with the experimental protocol by raising their arms during walking (n=2), and failure in identifying parts of the skeletons from the videos caused by the influence of loose black clothing (n=7).
Replication	The results can be replicated by following the methods described in the manuscript or using the code available in Code Ocean. A toy dataset including skeleton data, gait features extracted from the skeleton data, and the features extracted using the DMGrad-CAM method are provided to demonstrate the functionality of the code. However, the complete dataset, including clinical test scores and gait examination videos, cannot be released due to privacy concerns.
Randomization	The gait videos were randomly divided into training, testing, and validation datasets on a per-subject basis, ensuring that videos from the same individual were exclusively assigned to one dataset. This prevents videos of a single participant from being distributed across the training, testing, and validation sets.
Blinding	To compare the performance of clinicians and the deep learning model in assessing gait changes induced by medication, clinicians were blinded to the patient's medication status (i.e., with gait videos only) during one of the three evaluation rounds.

## 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 & 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
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

### 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 checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging