# nature portfolio

Corresponding author(s):	Rui Ponte Costa
Last updated by author(s):	Nov 30, 2022

## **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 <u>Editorial Policies</u> and the <u>Editorial Policy Checklist</u>.

٠.			. ~ +	100
_	_			ICS
_	LU	ı	IJL	

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
	The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
	The statistical test(s) used AND whether they are one- or two-sided Only common tests should be described solely by name; describe more complex techniques in the Methods section.
	A description of all covariates tested
	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
	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)
	For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i> ) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
X	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
X	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
	Estimates of effect sizes (e.g. Cohen's $d$ , Pearson's $r$ ), indicating how they were calculated
	Our web collection on <u>statistics for biologists</u> contains articles on many of the points above.

### Software and code

Policy information about <u>availability of computer code</u>

Data collection

Pytorch version 1.6.0 was used to implement the model. Examples of the code to run tasks are provided here (https://github.com/neuralml/ccDNI). Our deep learning implementation is based on that of (github.com/koz4k/dni-pytorch).

Data analysis

Demixed component analysis (https://github.com/machenslab/dPCA)

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 <u>availability of data</u>

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 have used standard machine learning data sets (freely available under a Creative Commons license): MNIST and COCO training data set ILSVRC-2012-CLS.

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.					
☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences					
For a reference copy of the document with all sections, see <a href="mailto:nature.com/documents/nr-reporting-summary-flat.pdf">nature.com/documents/nr-reporting-summary-flat.pdf</a>					
	nces study design close on these points even when the disclosure is negative.				
Sample size	To show robustness to initial conditions we used N=10 different initial conditions (seeds) in all the results we report. This is the order of magnitude commonly used in deep learning simulations. Our results show that this number of samples is sufficient to get a clear results between the proposed model and the control model. We tried using more samples, which did not change the results.				
Data exclusions	No data exclusions were conducted.				
Replication	Replication was conducted across different initial conditions of the model.				
Randomization	Randomization was conducted by randomly initializing the model across to provide different initial conditions of the model.				
Blinding	Not relevant as this is a theoretical study (i.e. no participants used).				

## 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		Methods	
n/a	Involved in the study	n/a Involved in the study	
$\boxtimes$	Antibodies	ChIP-seq	
$\boxtimes$	Eukaryotic cell lines	Flow cytometry	
$\boxtimes$	Palaeontology and archaeology	MRI-based neuroimaging	
$\boxtimes$	Animals and other organisms		
$\boxtimes$	Human research participants		
$\boxtimes$	Clinical data		
$\boxtimes$	Dual use research of concern		