nature portfolio

Corresponding author(s): Aydogan Ozcan

Last updated by author(s): Jun 17, 2023

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	
☐	☒	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. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	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 statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	The software and algorithms for collecting and processing experimental data are detailed in the manuscript, with references to published literature. Experimental data were collected using a custom-designed holographic microscope. Raw holograms (~4.6K × 3.5K pixels) were recorded by a complementary metal-oxide semiconductor (CMOS) with a pixel size of 1.12 μm (IMX 081 RGB, Sony Corp.) using AYA software tool (Sony Corp.) and LabView 2012 (version 12.0).
Data analysis	The codes for the deep learning models used in this work (written in Python 3.9.6 and PyTorch 1.9.0) can be accessed through: https://github.com/PORPHURA/GedankenNet The trained model and demo data are uploaded and available in the same code repository. The code for analyzing the results was written in Python using standard, open-source Python libraries.

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 involved simulation data generated from the COCO dataset (available at https://cocodataset.org/#home) and the artificial image dataset, and experimental data of human samples. Simulation data used in the manuscript can be generated using the open-source code of the manuscript, public datasets and open-source Python libraries. The training artificial image dataset is available at: https://github.com/PORPHURA/GedankenNet. A portion of the experimental dataset corresponding to human tissue samples is also shared and referenced in the code repository without any links or identifiers to the patients, which is made available at: https://github.com/PORPHURA/GedankenNet.
Additional data are available from the corresponding author upon reasonable request.

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 of the human tissue slides was not collected/unknown, not used in this manuscript, and was irrelevant to this research.
Population characteristics	All slides used in this manuscript were de-identified and prepared from existing specimens without links or identifiers to the patients. Therefore, this work did not interfere with standard practices of care or sample-collection procedures. The original tissue specimens had been archived before this work started, and had not been collected specifically for this research.
Recruitment	De-identified tissue slides were created before this work started and not collected specifically for this research.
Ethics oversight	Exempt: we used already existing, de-identified tissue samples collected before this work, without links or identifiers to the patients available.

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

Life sciences study design

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

Sample size	To prepare the multi-height experimental hologram training dataset presented in the manuscript, we collected ~100K input-target pairs of stained human lung, prostate, salivary gland, kidney, liver and esophagus tissue sections, Pap smears, and unstained label-free human kidney tissue sections, each of which is of 512*512 pixels. The size of the experimental hologram training dataset was chosen as ~100K to match with the size of the artificial image training dataset. Empirically 100K training images are sufficient to reach a balance between the reported models' training speed and performance. For the testing of the presented models, a subset of the lung, prostate, salivary gland slides from new patients and Pap smears were excluded from the training dataset and used as testing datasets, containing 94, 49, 49 and 47 unique 512*512 pixel field-of-views (FOVs), respectively. The holograms of the unstained (label-free) kidney tissue thin sections (~3-4 μm thick) were used as our phase-only object test dataset containing 98 unique FOVs. The size of testing datasets was empirically determined so that it is sufficient to demonstrate the distribution of evaluation metrics of interest.
Data exclusions	The training FOVs and testing FOVs were strictly exclusive, and the testing FOVs were captured strictly from different sample slides from the training slides. The ground truth complex fields recovered from raw CMOS holograms (~4.6K*3.5K) were cropped and only the central regions (~4K*3K) were used for training and testing because of the missing side diffraction patterns out of the CMOS region.
Replication	The presented framework was validated by training on three training datasets, including artificial holograms generated from random synthetic images, natural images (COCO) and experimental holograms of stained tissue sections. After training, all the models were blindly tested on multiple testing datasets, including experimental holograms of lung, prostate, salivary gland slides from new patients and Pap smears, containing 94, 49, 49, and 47 unique FOVs. The trained models were also blindly tested on artificial holograms generated from 100 testing random images and 100 testing COCO natural images. The models for phase-only objects were tested on the testing dataset of unstained, label-free human kidney tissue sections, containing 98 unique FOVs.

	The above model training and testing process was also repeated for four times with each selection of the number of input holograms (M). The ensemble models were evaluated and reported in Fig. 2.
Randomization	The training, validation and testing image datasets were randomly partitioned. The artificial images used for training were randomly generated using a Python package (randimage) as described in the manuscript. The training process randomly selected mini-batches for optimization.
Blinding	All the testing results of artificial and experimental holograms generated by the trained neural networks were blindly performed on new fields-of-view excluded from the training dataset. The blind testing experimental holograms were also captured on new slides from new patients that did not appear in the training dataset.

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
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging