

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 (<i>n</i>) 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 <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
☒	☐ 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>
☒	☐ 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 <i>d</i> , Pearson's <i>r</i>), 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 data collection code is included in an open source code repository.
Data analysis	The data analysis code is included in an open source code repository. https://github.com/GeoDS/Imagery2Flow

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](#)

All data used in this research are freely available to the public. The commuting flows data in the U.S is downloaded from the LODES program by the U.S Census Bureau 2020 survey results, accessible at <https://lehd.ces.census.gov/data/lodes/>. The geographic boundary of census tracts is accessible at <https://www2.census.gov/geo/tiger/TIGER2020/>. The Sentinel-2 Level 2-A, Landsat-8 Collection 2 and Landsat-5 Collection 2 multi-spectral optical pre-processed imagery are used in our study, with a spatial resolution of 10m x 10m and 30m x 30m, respectively. The satellite imagery is obtained and filtered for clouds and cloud-

shadows using Google Earth Engine. The travel distances between the centroids of census tracts (driving route in meters) are computed using the Open Source Routing Machine (OSRM). For auxiliary data, the land cover for 2020 and land use data for 2015 (the latest available) are sourced from the U.S. Geological Survey National Land Cover Database (NLCD) and HISDAC-US.

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	No human participants or biological material are used in this study.
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

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)

Behavioural & social sciences study design

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

Study description	In this work, we develop a graph-based deep learning framework Imagery2Flow for predicting fine-grained human mobility flows in urban areas using medium-resolution (10 meter Sentinel-2 to 30 meter Landsat) satellite imagery. Our experiments demonstrate good performance and flexible spatial-temporal generalizability on the top-10 largest Metropolitan Statistical Areas (MSAs) of the United States.
Research sample	The number of satellite imagery samples within census tracts in Metropolitan Statistical Areas (MSAs): 1. New York-Newark-Jersey City, NY-NJ-PA 4953 2. Los Angeles-Long Beach-Anaheim, CA 3112 3. Chicago-Naperville-Elgin, IL-IN-WI 2335 4. Dallas-Fort Worth-Arlington, TX 1704 5. Houston-The Woodlands-Sugar Land, TX 1606 6. Washington-Arlington-Alexandria, DC-VA-MD-WV 1487 7. Philadelphia-Camden-Wilmington, PA-NJ-DE-MD 1556 8. Miami-Fort Lauderdale-Pompano Beach, FL 1497 9. Atlanta-Sandy Springs-Alpharetta, GA 1500 10. Phoenix-Mesa-Chandler, AZ 1104
Sampling strategy	In each MSA, stratified sampling is conducted based on population in each area, with census tracts divided into a 60% training set, 20% validation set, and 20% test set. The 5-fold cross-validation is utilized to get the human mobility flow prediction covering the entire MSA.
Data collection	This study is a secondary analysis of publicly available, aggregated survey data. The data were collected and processed by the U.S. Census Bureau. No direct interaction with participants was performed by the authors. The commuting dataset used in this study was collected from LODES between 2002 to 2020. Data are aggregated and released on a yearly basis.
Timing	The year of 2020 Sentinel-2 to Landsat satellite imagery were used for training while the 2010 and 2002 data were used for temporal transferability testing.
Data exclusions	We removed the flows with fewer than 10 commuting workers between census tracts due to their insignificance in accurately capturing stable movement patterns between census tracts.
Non-participation	No participants in this study
Randomization	No participants in this study

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
☒	☐ Plants

Methods

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

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A