Stanford University
Human-Centered
Artificial Intelligence

Policy Brief
HAI Policy & Society
July 2024

Using AI to Map Urban Change

Tianyuan Huang, Zejia Wu, Jiajun Wu, Jackelyn Hwang, and Ram Rajagopal

CITIES ARE CONSTANTLY EVOLVING, AND BETTER UNDERSTANDING THOSE CHANGES facilitates better urban planning and infrastructure assessments and leads to more sustainable social and environmental interventions. Researchers currently use data such as satellite imagery to study changing urban environments and what those changes mean for public policy and urban design. But flaws in the current approaches, such as inadequately granular data, limit their scalability and their potential to inform public policy across social, political, economic, and environmental issues.

Street-level images offer an alternative source of insights. These images are frequently updated and high-resolution. They also directly capture what's happening on a street level in a neighborhood or across a city. Analyzing street-level images has already proven useful to researchers studying socioeconomic attributes and neighborhood gentrification, both of which are essential pieces of information in urban design, sustainability efforts, and public policy decision-making for cities. Yet, much like other data sources, street-level images present challenges: accessibility limits, shadow and lighting issues, and difficulties scaling up analysis.

To address these challenges, our paper "CityPulse: Fine-Grained Assessment of Urban Change with Street View Time Series" introduces a multicity dataset of labeled street-view images and proposes a novel

Key Takeaways

Physical changes in cities offer profound insights into demographic shifts, housing trends, neighborhood gentrification, and other issues critical to urban planning and sustainability. But many current methods for analyzing urban change are limited as they rely heavily on surveys and construction permit data.

We use Google Street View data from across the United States to curate the largest street-view time series dataset to date and build an AI model to evaluate changes in the images. Our time series data allows the AI model to more easily filter out irrelevant variations and instead focus on detecting large-scale changes to urban environments.

Using a case study in Seattle, we demonstrate that our method for urban change detection provides a scalable, more accurate proxy for assessing neighborhood and citywide socioeconomic changes.

Policymakers should look to street-view image data and AI models to identify areas undergoing change with more precision, which can enable more efficient resource allocation and earlier intervention to mitigate the potential negative impacts of gentrification.

artificial intelligence (AI) model to detect urban changes such as gentrification. We demonstrate the change-detection model's effectiveness by testing it on images from Seattle, Washington, and show that it can provide important insights into urban changes over time and at scale. Our data-driven approach has the potential to allow researchers and public policy analysts to automate and scale up their analysis of neighborhood and citywide socioeconomic change.

Physical changes in urban environments offer profound insights into urban policies and economics.

Introduction

Physical changes in urban environments offer profound insights into urban policies and economics, including housing value trends and spatial segregation effects. They also provide insights into downstream impacts, such as neighborhood gentrification and disaster recovery. At present, researchers predominantly use datasets made up of satellite and aerial imagery, survey data, and building permit data to measure physical urban change, but these methods have many limitations. Survey data is often inadequately granular, both spatially and temporally; top-down views from remote sensing data incompletely capture street-level changes that impact daily life in cities; and data on construction permits is not always accessible.

Street-view imagery provides a more comprehensive spatial representation of city development over time. Such imagery has already been used to quantify urban greenery, indicate region functions (e.g., areas used for public health or recreational purposes), uncover economic and sociodemographic patterns, predict a population's well-being, and estimate buildings' energy efficiency. Imagery of city streets can also be used to broadly evaluate how buildings, streets, and other city elements deteriorate and change. But current methods rely on comparing pairs of historical street-views for individual locations, which doesn't capture the full range of transformations that can occur in cities. A comprehensive time series of street-view data would enable a more granular, insightful, and generalizable analysis of urban environments.

With that objective in mind, we collected and curated a Google Street View time series dataset with images on about 1,000 coordinates in six cities, collected from 2007 to 2023—the largest street-view change-detection dataset available to date. Each time series contains around 10 images over the 16-year period, and each image is labeled with "change" or "no change." We then leveraged past work on the use of neural networks and machine learning techniques for change detection, commonly used in remote sensing to identify differences between two images, to train an AI model on our street-view time series dataset to analyze urban change. The AI model can extract object details and structural information about the built environment and mitigate noisy effects, thereby enabling a more fine-grained analysis of urban change.

It also enables change detection at scale—so researchers can ingest more data and analyze more cities.

Research Outcomes

We ran a series of tests and comparisons on 50 percent of the street-view images in our dataset taken from 25 locations in Seattle, 13 locations in San Francisco, 21 locations in Oakland, 97 locations in Los Angeles, and 29 locations in Boston—for a total of 12,221 image pairs. The tests had three objectives: 1) Evaluate how "backbone" models (i.e., models that extract image features) perform when doing change detection on street-view images; 2) Experiment with and assess the advantages of using street-view time series data; and 3) Compare how two different self-supervised pre-training methods (i.e., different ways of training models on unlabeled datasets, whereby the models generate the labels themselves) perform on our change-detection tasks for the image data.

The AI model can extract object details and structural information about the built environment and mitigate noisy effects, thereby enabling more fine-grained analysis of urban change.

First, Meta AI's DINOv2, a self-supervised model for image-level and pixel-level analysis, demonstrated the best backbone AI model performance of the four pre-trained visual models we tested. With 88.85 percent accuracy after fine-tuning, it handled shadows, obstructions in images, and other changes almost exactly as well as a human analyst, who labeled images with 90 percent accuracy. DINOv2 evidently has the capacity to identify visual features that are well-suited for change detection.

Second, we found that our change-detection model performs much more accurately when analyzing our comprehensive street-view time series data than when comparing only pairs of street images. The time series data gives the model sufficient information to identify and filter out temporary changes in lighting, vegetation, and vehicles as well as other irrelevant variations that occur over time and which are all too common in an urban environment. This allows the model to focus on detecting large-scale changes in the urban environment such as building constructions.

And third, we compared two approaches to pre-training AI models for street-image analysis: StreetMAE, which randomly masks and then reconstructs patches in images; and StreetBYOL, which uses two neural networks that interact with and learn from each other. StreetBYOL performed much better, likely because it has an added feature that labels every pixel with a word (also known as "semantic segmentation") and because StreetMAE's reconstruction of pixels is more prone to reproduce the noise in the original images. Nonetheless, DINOv2 still outperformed both of these domain-specific pre-trained models.

Finally, we conducted a case study on urban change in Seattle to evaluate how generalizable our change-detection model is on a large scale. We ran the model on 795,919 images from 118,592 locations, collected between 2007 and 2023. After preparing census tract-level data from the American Community Survey on population size and median household incomes, we correlated those variables against an analysis of the time series data and an analysis of a construction permit dataset we obtained. The time series data yielded a significant correlation with changes in population size and median household income while the construction permit data did not. This highlights that our proposed model for detecting urban change can complement existing indicators of urban change.

Street-image data and AI models are driving new ways for researchers, policy analysts, and urban designers to deeply understand how cities evolve over time.

Policy Discussion

Analyzing physical changes in urban environments allows policymakers to gain essential insights into demographic shifts, neighborhood gentrification, urban greenery, disaster recovery, buildings' energy efficiency, the uses of public spaces, and much more. Today, street-image data and AI models are driving new ways for researchers, policy analysts, and urban designers to deeply understand how cities evolve over time and the resulting social, spatial, and economic impacts.

Policymakers should utilize our proposed urban change analysis method to better inform urban policy using street-view data and AI capabilities. Our curated dataset is the largest street-level scene change-detection dataset currently available. Combined with our end-to-end change-detection pipeline that uses AI models to quickly track high-definition, on-the-ground visual proxies of urban change, our paper lays a foundation for additional work that improves model performance on these tasks. The prior heavy reliance on survey data (e.g., from the American Community Survey)—which lacks spatial and temporal granularity—for change-detection methods also underscores the need for greater curation of and access to street-view image data.

Our change-detection approach could significantly improve the precision of urban change analysis for policy institutions, such as city-level urban planning offices and federal housing or environmental authorities. Models such as Meta AI's DINOv2, which have already approached a human-level ability to detect urban change, can supplement and improve human analysis of urban image data. For example, they allow us to more accurately detect gentrification and—when combined with additional data—gain a more granular understanding of its consequences, enabling more effective allocation of resources and earlier interventions. More broadly, insights gleaned from these models could inform our understanding of barriers to and progress toward achieving sustainable development goals—such as advancing inclusivity or mapping different neighborhoods' access to and use of public spaces such as parks.

Importantly, policymakers should recognize that this approach has limitations. Street-view data focuses on changes that are observable from the street level outside, excluding interior renovations. Google Street View data is also not evenly distributed across space and time, and Google has been less consistent in updating street-image data in many developing countries and regions outside the United States. Policymakers may consider identifying opportunities to improve public access to government data and increase opportunities for collaborative public-private research on urban change, such as setting up local research projects between cities and academic institutions.

Our novel approach to urban change analysis enables researchers to quickly, accurately, and regularly measure demographic shifts, neighborhood gentrification, and other issues of great interest to policymakers and society.

[These models] allow us to more accurately detect gentrification and [...] gain a more granular understanding of its consequences, enabling more effective allocation of resources and earlier interventions.

Reference: The original article is accessible at Tianyuan Huang et al., “**CityPulse: Fine-Grained Assessment of Urban Change with Street View Time Series**,” arxiv.org, January 3, 2024, https://arxiv.org/abs/2401.01107.

Stanford University’s Institute for Human-Centered Artificial Intelligence (HAI) applies rigorous analysis and research to pressing policy questions on artificial intelligence. A pillar of HAI is to inform policymakers, industry leaders, and civil society by disseminating scholarship to a wide audience. HAI is a nonpartisan research institute, representing a range of voices. The views expressed in this policy brief reflect the views of the authors. For further information, please contact **HAI-Policy@stanford.edu**.

Tianyuan Huang is a PhD student in civil and environmental engineering at Stanford University.

Zejia Wu is a PhD student in electrical and computer engineering at the University of California San Diego.

Jiajun Wu is an assistant professor of computer science at Stanford University and a faculty affiliate of the Stanford Institute for Human-Centered Artificial Intelligence (HAI).

Jackelyn Hwang is an assistant professor of sociology at Stanford University, director of the Changing Cities Research Lab, and a faculty affiliate of Stanford HAI.

Ram Rajagopal is an associate professor of civil and environmental engineering and of electrical engineering at Stanford University and a faculty affiliate of Stanford HAI.

Stanford University
Human-Centered
Artificial Intelligence

Stanford HAI: 353 Jane Stanford Way, Stanford CA 94305-5008
T 650.725.4537 **F** 650.123.4567 **E** HAI-Policy@stanford.edu **hai.stanford.edu**