scientific **data**

OPEN

DATA DESCRIPTOR

A Satellite Imagery Dataset for Long-Term Sustainable Development in United States Cities

Yanxin Xi [1], Yu Liu [2,3], Tong Li [2,3], Jingtao Ding [2,3], Yunke Zhang [2,3], Sasu Tarkoma [1], Yong Li [2,3] ✉ & Pan Hui [1,4,5] ✉

Cities play an important role in achieving sustainable development goals (SDGs) to promote economic growth and meet social needs. Especially satellite imagery is a potential data source for studying sustainable urban development. However, a comprehensive dataset in the United States (U.S.) covering multiple cities, multiple years, multiple scales, and multiple indicators for SDG monitoring is lacking. To support the research on SDGs in U.S. cities, we develop a satellite imagery dataset using deep learning models for five SDGs containing 25 sustainable development indicators. The proposed dataset covers the 100 most populated U.S. cities and corresponding Census Block Groups from 2014 to 2023. Specifically, we collect satellite imagery and identify objects with state-of-the-art object detection and semantic segmentation models to observe cities' bird's-eye view. We further gather population, nighttime light, survey, and built environment data to depict SDGs regarding poverty, health, education, inequality, and living environment. We anticipate the dataset to help urban policymakers and researchers to advance SDGs-related studies, especially applying satellite imagery to monitor long-term and multi-scale SDGs in cities.

Background & Summary

Nowadays, more than 50% of the population lives in cities, producing 80% of the GDP worldwide[1,2]. Therefore, cities play an increasingly important role in achieving United Nations Sustainable Development Goals[3] (SDGs), which aim to prosper economic growth and meet social needs. According to the report "The United States Sustainable Development Report[4]", cities in the United States (U.S.) perform poorly on a series of SDGs (e.g., Boise city lags behind in quality education, and Raleigh city shows high poverty rate[5]). Currently, monitoring sustainable development in U.S. cities heavily relies on door-to-door surveys such as the American Community Survey[6,7] (ACS) data. First, ACS data for constructing the SDG index in U.S. cities is economically costly as the annual budget can reach millions of dollars[8]. Second, the current SDG index dataset for U.S. cities is meant for a single year and only focuses on the city level, which hinders monitoring of multi-scale and multi-year SDG progress[9]. Alternatively, built upon the rapid development of remote sensing and deep learning techniques, satellite imagery showing nearly real-time and bird's-eye view information in cities has been broadly investigated as a data source for SDG monitoring[10–17]. Therefore, monitoring SDGs in cities with satellite imagery is of great significance in promoting sustainable urban development.

However, a long-term and multi-scale satellite imagery dataset, which reveals the yearly change in SDGs in multiple years and in different spatial (administrative) scales for city SDG monitoring, is still lacking. For instance, some satellite imagery datasets about SDGs focus on either country level or cluster level (25–30 households) spatially and only contain data of a single year[18], which barely match the requirements of long-term and multi-scale SDG monitoring in cities. Other open-source satellite imagery datasets, such as SpaceNet[19]

[1]Department of Computer Science, University of Helsinki, Helsinki, Finland. [2]Beijing National Research Center for Information Science and Technology (BNRist), Beijing, P. R. China. [3]Department of Electronic Engineering, Tsinghua University, Beijing, P. R. China. [4]Computational Media and Arts Thrust, Hong Kong University of Science and Technology (Guangzhou), Guangzhou, P. R. China. [5]Division of Emerging Interdisciplinary Areas, Hong Kong University of Science and Technology, Hong Kong, P. R. China. ✉e-mail: liyong07@tsinghua.edu.cn; panhui@ust.hk
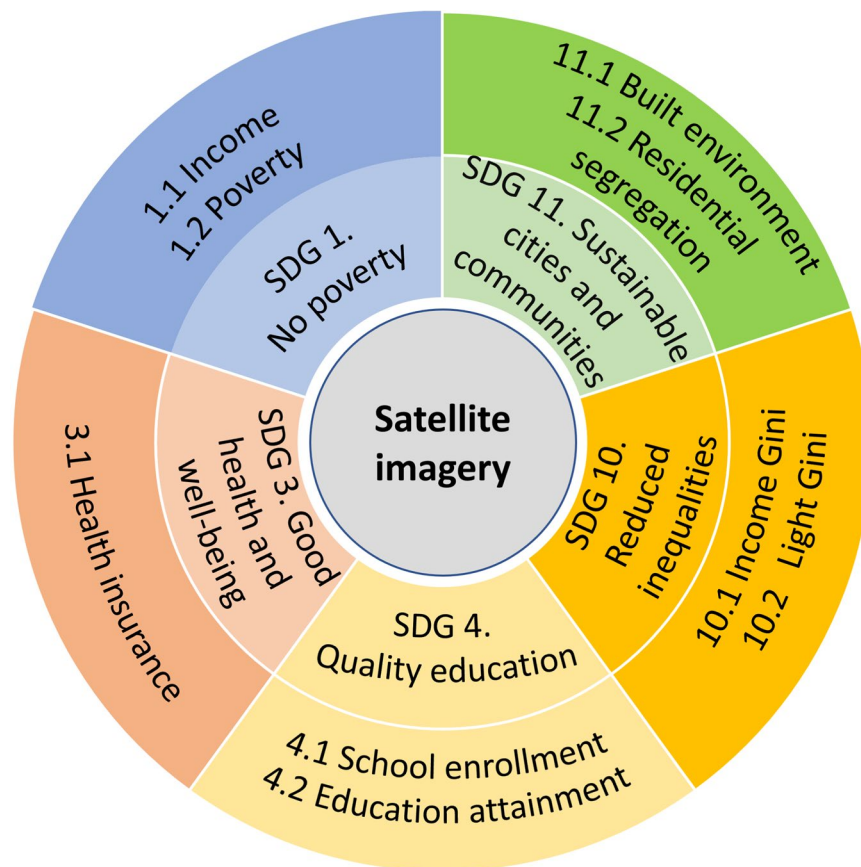
1

Fig. 1 Schematic overview of the target dataset.

or ForestNet[20], merely contain the dataset for one single SDG. Besides, other survey data for SDG, such as UNESCO Survey on Public Access to Information and Survey Module[21] on SDG Indicator 16.b.1 & 10.3.1, are based on questionnaires. At last, although plenty of survey data in the U.S. will aid in SDG monitoring, it would be difficult for urban policymakers and researchers to extract the critical information easily. Motivated by the SustainBench[22] in low-income countries, and to fulfill the data requirements for SDG monitoring in U.S. cities, we propose a comprehensive long-term and multi-scale satellite imagery dataset with 25 SDG indicators for five SDGs (SDG 1, SDG 3, SDG 4, SDG 10, and SDG 11). Moreover, the dataset covers about 45,000 Census Block Groups (CBGs) in the 100 most populated U.S. cities from 2014 to 2023. Using satellite images and SDG data in the U.S., urban policymakers and researchers can develop various models or assumptions regarding SDG monitoring remotely. And further, the dataset from the U.S. can aid urban policymakers and researchers in inferring SDG progress in low-income countries, which mostly lack SDG-related survey data.

Figure 1 presents the scheme of our produced dataset with two components: the satellite imagery data containing the detected objects and land cover semantics obtained with state-of-the-art deep learning models and the corresponding SDG indicators in 100 cities from 2014 to 2023. For the satellite imagery data, we incorporate the daytime satellite imagery with the spatial resolution of 0.3 m and several objects (such as truck and basketball court[23–26]) detected as well as several land cover semantics (forest and road[27]) inferred with the models transferred from the computer vision community. The detected objects refer to the countable artificial objects and venues visible in satellite images in cities, while the land cover is mostly the uncountable environmental information. For SDGs, we collect the indicators that can be inferred from satellite images in urban scenarios. Specifically, indicators for SDG 1 "No poverty", SDG 3 "Good health and well-being", SDG 4 "Quality education", SDG 10 "Reduced inequalities", and SDG 11 "Sustainable cities and communities"[3,28] are included in our dataset. The indicators are generated from multi-source data, including nighttime light (NTL) data from Earth Observation Group (EOG)[29], WorldPop population data[30], ASC data[6,7], and OpenStreetMap (OSM) built infrastructure[31,32].

At last, this paper advances the SDG-related community by generating a long-term and multi-scale satellite imagery dataset in urban scenarios by collecting and processing satellite images and SDG indicators from multiple sources, which is a time-consuming and laborious work, and the alignment of satellite image visual attributes and SDG data. The dataset aims to help urban policymakers and researchers, who might not have the platform to collect and process the large volume of data, to conduct numerous SDG studies spanning poverty, health, education, inequality, and built environment. More importantly, as the first urban satellite imagery dataset for multiple SDGs monitoring with interpretable visual attributes (e.g., cars, buildings, roads, etc.), it can aid in further achieving sustainable cities with high interpretability and advancing urban sustainability progress. The
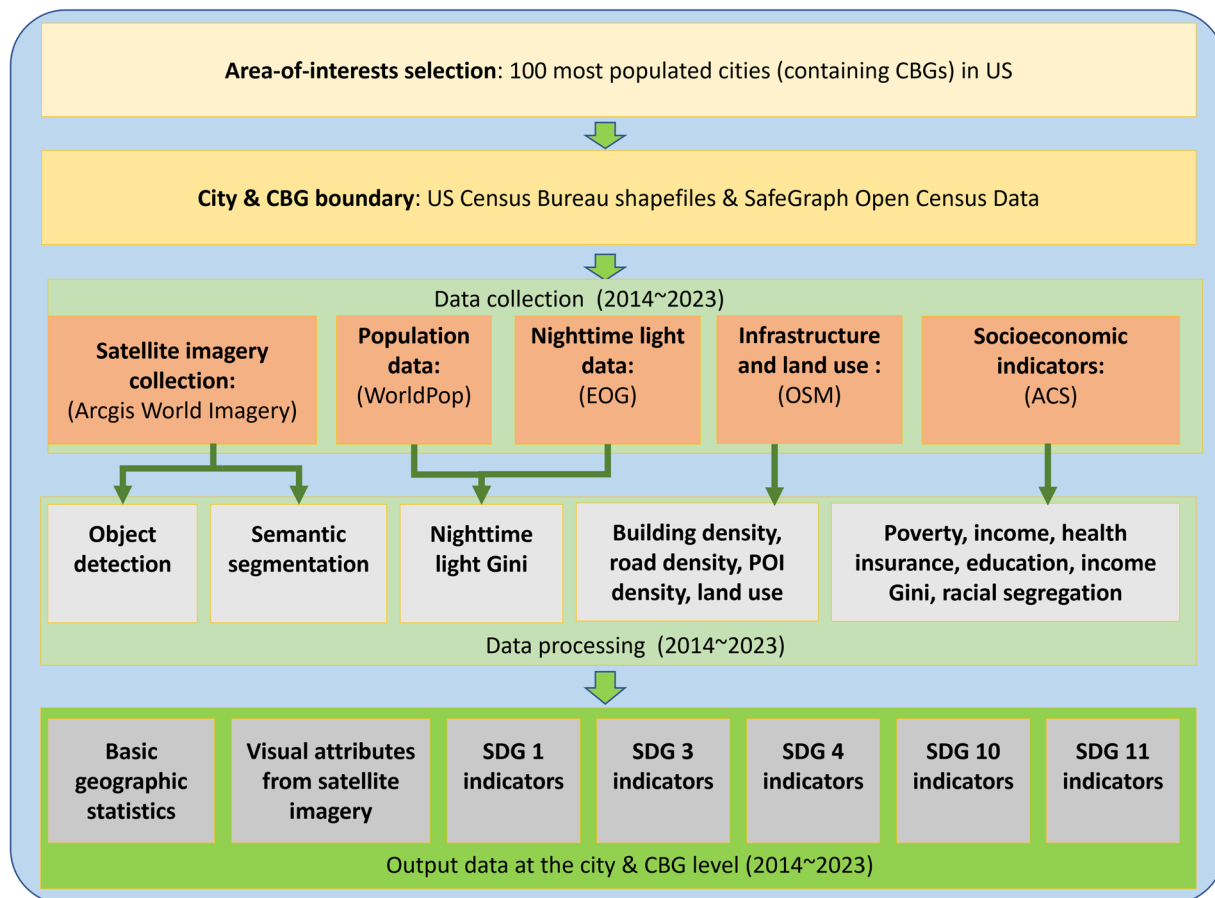
Fig. 2 Overall workflow of the dataset generation process.

satellite imagery and the visual attributes extracted by the computer vision models in the dataset can serve as the input for various kinds of research regarding SDGs, and the SDG indicators act as output. Specifically, we recommend the following potential applications:

- Researchers can design deep learning models to predict various long-term SDGs (income, poverty, and built environment) in cities from historical satellite images.
- Researchers can also estimate various SDG progresses by utilizing multi-scale satellite imagery visual attributes at the CBG and city levels and reveal the linkage between the multi-scale satellite imagery and SDGs.
- Researchers can propose a spatiotemporal framework that simultaneously utilizes long-term and multi-scale satellite imagery for SDG monitoring, which sheds light on satellite imagery fusion of temporal and spatial dimensions.

Methods

We aim to provide a comprehensive and representative dataset that includes satellite imagery and corresponding SDG indicators covering long terms and multiple scales. To ensure that the indicators can thoroughly depict sustainable urban development, we select five SDGs altogether: SDG 1 No poverty (five indicators), SDG 3 Good health and well-being (five indicators), SDG 4 Quality education (five indicators), SDG 10 Reduced inequalities (two indicators), and SDG 11 Sustainable cities and communities (eight indicators). Overall, the target dataset generation process includes collecting, processing, and aligning multi-source data, and the overall workflow is presented in Fig. 2. We first select the 100 most populated cities and gather the corresponding CBG/city boundaries. Second, we collect satellite imagery, population, NTL, OSM, and ACS data from multiple sources. At last, we process the multi-source data and produce the final output data at the CBG and city levels, containing basic geographic statistics, satellite imagery attributes, and SDG indicators.

Determining the area-of-interests and boundary extraction. We select the 100 cities with the most population in the contiguous United States, which is explored on the ACS 2021 population data[33]. The population in the 100 cities varies from 222,194 to 8,467,513, with a mean population of 642,002. The city-of-interests and population in descending order are shown in Table 1.

City	Population	City	Population	City	Population	City	Population
New York	8,467,513	Los Angeles	3,849,306	Chicago	2,696,561	Houston	2,287,047
Phoenix	1,624,539	Philadelphia	1,576,251	San Antonio	1,451,863	San Diego	1,381,600
Dallas	1,288,441	San Jose	983,530	Austin	964,000	Jacksonville	954,624
Fort Worth	940,437	Columbus	907,310	Indianapolis	882,327	Charlotte	879,697
San Francisco	815,201	Seattle	733,904	Denver	711,463	Oklahoma	687,691
Nashville-Davidson	678,845	El Paso	678,422	Washington	670,050	Boston	654,281
Las Vegas	646,776	Portland	642,218	Detroit	632,589	Louisvill -Jefferson	628,577
Memphis	628,118	Baltimore	576,498	Milwaukee	569,326	Albuquerque	562,591
Fresno	544,500	Tucson	543,215	Sacramento	525,028	Mesa	509,492
Kansas	508,415	Atlanta	496,480	Omaha	487,299	Colorado Springs	483,969
Raleigh	469,502	VirGinia Beach	457,672	Long Beach	456,063	Miami	439,906
Oakland	433,797	Minneapolis	425,338	Tulsa	411,905	Bakersfield	407,581
Wichita	395,707	Arlington	392,802	Aurora	389,675	Tampa	387,037
New Orleans	376,971	Cleveland	368,006	Anaheim	345,935	Henderson	322,202
Stockton	322,107	Lexington -Fayette	321,793	Corpus Christi	317,768	Riverside	317,257
Santa Ana	309,468	Orlando	309,193	Irvine	309,014	Cincinnati	308,913
Newark	307,216	St. Paul	307,176	Pittsburgh	300,454	Greensboro	298,250
St. Louis	293,310	Lincoln	292,648	Plano	287,037	Durham	285,439
Jersey	283,943	Chandler	279,445	Chula Vista	277,211	Buffalo	276,804
North Las Vegas	274,146	Gilbert	273,138	Madison	269,162	Reno	268,843
Toledo	268,504	Fort Wayne	263,814	Lubbock	260,990	St. Petersburg	258,214
Laredo	258,014	Irving	254,190	Chesapeake	251,269	Winston-Salem	250,337
Glendale	249,627	Enterprise CDP	245,286	Scottsdale	242,754	Garland	241,870
Boise	237,457	Norfolk	235,089	Arlington CDP	232,965	Spokane	229,065
Fremont	227,523	Richmond	226,604	Santa Clarita	224,588	San Bernardino	222,194

Table 1. City-of-interests and corresponding population in our dataset.

City Name	City GeoID	Track Code	Block Group	CBG Code
New York	1600000US3651000	17500	2	360050175002
New York	1600000US3651000	14100	1	360050141001
New York	1600000US3651000	14500	2	360050145001
…				
Los Angeles	1600000US0644000	216401	1	060372164011
Los Angeles	1600000US0644000	216401	2	060372164012
Los Angeles	1600000US0644000	216402	1	060372164021
…				

Table 2. Example of the geographic lookup table between cities and CBGs in the produced dataset.

Then we collect city geographic boundary files from the U.S. Census Bureau TIGER/Line shapefiles[34]. The shapefiles are divided by states, and each shapefile contains the city name (called "place" in the file), state name, Federal Information Processing Standard state code, and geographical boundary coordinates. We use the python packet *shapely* to access the shapefiles and extract the boundary coordinates using the city and state names. The geographic coordinate system is WGS84.

Next, we determine the corresponding CBGs within the cities. The boundaries of all CBGs in the U.S. are gathered from SafeGraph Open Census Data[7]. The CBG boundaries for the years 2014~2019 are the same, and the U.S. government adjusts the CBG boundary for the year 2020.

For each city, we overlap the CBG boundaries on the city boundary, and every CBG whose area intersection with city boundary takes up more than 10% of the corresponding CBG area is considered contained in the city. This process uses Python packets *shapely* and *geopandas*. The geographic lookup table between cities and corresponding CBGs is shown in Table 2. Till this step, we have the selected 100 most populated cities and corresponding CBGs spatially contained as the area-of-interests in our target dataset.

Processing of satellite imagery. Satellite imagery provides a near real-time bird's-eye view of the earth's surface. Combined with machine learning techniques, satellite imagery has been widely used in predicting socio-economic status, especially in urban research, which includes poverty/asset prediction[11,14,17], urban pattern mining[15], commercial activity prediction[16,35], and population prediction[12,36]. Inspired by the interpretable feature generation from satellite imagery[14], we provide satellite imagery visual attributes in our dataset to promote the

Objects Detected	passenger vehicles, swimming pools, planes, airports, trucks, railway vehicles, ships, engineering vehicles, bridges, roundabouts, vehicle lots, soccer fields, basketball courts, ground track fields, baseball diamonds, tennis courts, buildings (number of buildings)
Land Cover Semantics	road, forest, building (pixel percentage), water, barren, agriculture, background

Table 3. Visual attributes extracted from satellite images.

research of SDG monitoring. The processing of satellite imagery consists of three parts: imagery collection, object detection, and semantic segmentation.

First, we collect the satellite imagery in our dataset from Esri World Imagery[37]. It provides users access to the World Imagery of different versions created over time. The imagery is in RGB format collected from different satellites and of different spatial resolutions marked by different zoom levels, which split the entire world into different numbers of tiles. Overall, the imagery collection process includes generating image tile numbers according to the boundary of each city as well as the desired zoom level (spatial resolution) and downloading images with the tile numbers from the satellite imagery archive. In our target dataset, we set the zoom level to 19, which is about 0.3 m/pixel. We also select the Esri World Imagery archive of June from 2014 to 2023 to collect the satellite images of the 100 most populated cities, which generates altogether 12,269,976 images each year.

Second, many aspects of cities are related to people's lives and can reveal SDG progress. Transportation in the city is integral to urban development[38], and further, transportation and mobility were recognized as central to sustainable development at the 2012 United Nations Conference on Sustainable Development[39]. Sports & leisure are highly correlated to citizens' life quality[40,41]. Children and young people benefit largely from sports, which are inseparable from a quality school education, promoting SDG 3 and SDG 4[42]. The building characteristics (e.g., building type) can reveal the population and income status in urban areas[43,44], and the impact of buildings on human well-being can not be neglected. Therefore, the buildings, cars, and other objects in satellite imagery contain certain correlations with SDG indicators. In our dataset, we consider 17 objects from the abovementioned aspects: transportation, sports & leisure, and building.

The urban object categories are presented in Table 3. We use the YOLOv5s model[45,46] pre-trained on the MS COCO dataset[47] and finetune it on xView dataset[23] and DOTA v2 dataset[25] to detect objects in the collected satellite imagery. The default parameters[48] are used for finetuning the object detection models. We aggregate the number of objects detected from satellite images at the CBG and city levels to show visual object attributes at multiple scales.

Third, land cover information such as forests or water can also depict the urban environment and is not included in the detected objects. Therefore, we add the land cover semantic information inferred from satellite imagery in our generated dataset. We use the Vision Transformer (ViT)-Adapter-based semantic segmentation model[49–51] pre-trained on the ADE20K dataset[52] and finetune it on LoveDA dataset[27] to generate semantic information from the collected satellite imagery, which includes background, building, road, water, barren, forest, and agriculture. Moreover, we compute the pixel-level percentage of each semantic information presented in Table 3 in each satellite imagery and aggregate them at the CBG and city levels, respectively.

Processing of basic geographic statistics. For each CBG/city, we present the population, area, centroid coordinates, and geographic boundary, which describe the essential information for the selected area-of-interests. Specifically, we collect the population data from 2014 to 2020 from the WorldPop project[30,53]. The population data is downloaded at a resolution of 3 arc (approximately 100 m at the Equator). We use Python packets *shapely* and *gdal* to crop the population data with the CBG/city geographic boundary and sum up the cropped pixel values as the total population. The area (km^2) is calculated from the CBG/city boundary data with Python packet *geopandas*. The geographic centroid can also be computed with Python packet *geopandas*.

Processing of SDG indicators. There are five SDGs (SDG 1, SDG 3, SDG 4, SDG 10, and SDG 11) concerning poverty, health, education, inequality, and built environment collected in our produced dataset at the CBG/city level. SDG 1 "No poverty" focuses on income and population in poverty status. The indicators for "No poverty" are collected from ACS data. SDG 3 "Good health and well-being" and SDG 4 "Quality education" highlight people's health insurance status and population with different academic degrees, and corresponding indicators are extracted from ACS data. SDG 10 "Reduced inequalities" intends to reduce inequality, and the indicators are from ACS data and from NTL combined with population data with a recent algorithm for monitoring regional inequality through NTL[54]. Finally, SDG 11 "Sustainable cities and communities" reflects the living conditions in CBG/city, and the related indicators are calculated from OSM historical data and ACS data. Altogether, we collect 25 indicators across five SDGs. The indicators and relevant SDG targets are described in Table 4. Specifically, there are eight SDG targets included in this dataset:

- Target 1.2: By 2030, reduce at least by half the proportion of men, women, and children of all ages living in poverty in all its dimensions according to national definitions.
- Target 1.4: By 2030, ensure that all men and women, in particular the poor and the vulnerable, have equal rights to economic resources, as well as access to basic services, ownership and control over land and other forms of property, inheritance, natural resources, appropriate new technology and financial services, including microfinance.
- Target 3.8: Achieve universal health coverage, including financial risk protection, access to quality essential healthcare services, and access to safe, effective, quality, and affordable essential medicines and vaccines for all.

SDG Target	Indicator	SDG Target	Indicator
1.2	**Population Below Poverty**	1.2	**Population Above Poverty**
1.2	**Population With A Ratio Of Income To Poverty Level Under 0.5**	1.2	**Population With A Ratio Of Income To Poverty Level 0.5 to 0.99**
1.4	**Median Household Income**		
3.8	**Population With No Health Insurance Under 18**	3.8	**Population With No Health Insurance Between 18 To 34**
3.8	**Population With No Health Insurance Between 35 To 64**	3.8	**Population With No Health Insurance Over 65 Years Old**
3.8	**Civilian Noninstitutionalized Population**		
4.1	**Population Enrolled In College**	4.3	**Population That Graduated From High School**
4.3	**Population With A Bachelor's Degree**	4.3	**Population With A Master's Degree**
4.3	**Population With A Doctorate**		
10.2	**Light Gini**	10.2	**Income Gini**
11.2	**Driving Road Density**	11.2	**Cycling Road Density**
11.2	**Walking Road Density**	11.3	**POI Density**
11.3	**Building Density**	11.3	**Land Use**
11.3	**Entropy Index**	11.3	**Index of Dissimilarity**

Table 4. Relationship between the SDG targets and the selected indicators.

- Target 4.1: By 2030, ensure that all girls and boys complete free, equitable, and quality primary and secondary education leading to relevant and effective learning outcomes.
- Target 4.3: By 2030, ensure equal access for all women and men to affordable and quality technical, vocational and tertiary education, including university.
- Target 10.2: By 2030, empower and promote the social, economic and political inclusion of all, irrespective of age, sex, disability, race, ethnicity, origin, religion or economic or other status.
- Target 11.2: By 2030, provide access to safe, affordable, accessible and sustainable transport systems for all, improving road safety, notably by expanding public transport, with special attention to the needs of those in vulnerable situations, women, children, persons with disabilities and older persons.
- Target 11.3: By 2030, enhance inclusive and sustainable urbanization and capacity for participatory, integrated and sustainable human settlement planning and management in all countries.

Indicators for SDG 1 "No poverty". SDG 1 aims to end poverty in all its forms everywhere[3]. Our target dataset incorporates income and poverty status data to represent the SDG 1 indicators in cities. Specifically, median household income, population above poverty (number of population whose income in the past 12 months is at or above poverty level), population below poverty (number of population whose income in the past 12 months is below poverty level), and population with a ratio of income to poverty level (the total income divided by poverty level) under 0.5 and between 0.5 to 0.99 are collected to describe the income & poverty in CBG/city. The poverty threshold is computed by the Census Bureau according to the family size and ages of family members every year with variations to Consumer Price Index. The threshold is a country-specific value and does not change geographically[55]. Moreover, population above/below poverty and population with different ratios of income to poverty level are measurements of poverty status.

We collect the median household income, population above/below poverty, and population with a ratio of income to poverty level under 0.5 and between 0.5 to 0.99 at the CBG level from the ACS data[6,7,56]. Then, we generate the city-level indicators: population above/below poverty and population with a ratio of income to poverty level under 0.5 and between 0.5 to 0.99 by aggregating all the CBG data within the city. Median household income at the city level is related to the income distribution of the population in cities and is gathered directly from ACS data[57]. The boundary files and ACS data are both collected from the U.S. Census Bureau. And ACS data denotes the city as "place" as in the boundary files, and the ACS definition of a city boundary is the same as the U.S. Census Bureau TIGER/Line shapefiles.

Indicators for SDG 3 "Good health and well-being". SDG 3 aims to ensure healthy lives and promote well-being for all populations at all ages[3]. In our target dataset, we use the population data with no health insurance covering all ages to represent SDG 3 indicators because health insurance is correlated to the health status of the population in urban regions[58,59]. Specifically, civilian noninstitutionalized population, population with no health insurance under 18, between 18 to 34, between 35 to 64, and over 65 years old are collected from ACS data[7] to describe the health insurance at the CBG and city levels.

Indicators for SDG 4 "Quality education". SDG 4 aims to ensure inclusive and equitable quality education and promote lifelong learning opportunities for all[3]. Therefore, indicators directly depicting city education status can be selected here. In dataset generation, we collect from ACS data[7] population enrolled in college, population that graduated from high school, population with a bachelor's degree, a master's degree, and a doctorate for indicators of school enrollment & education attainment to monitor SDG 4.

Indicators for SDG 10 "Reduced inequalities". SDG 10 aims to reduce inequality within and among countries[3]. We use income Gini[60] and light Gini[54] to monitor the process of SDG 10. The income Gini reveals the inequality status of income and is collected from ACS data. Light Gini can present the distribution of NTL per person and thus indirectly reveal regional development inequality. Similar to the income Gini, the lower the light Gini is, the more equally the region develops, which means the region moves towards eliminating inequality in SDG 10. The results in the original paper[54] report the light Gini at a 1-degree grid cell, which can not be directly used in urban scenarios. Therefore, we calculate the light Gini following the method[54]. Specifically, the NTL per person is calculated by dividing the NTL value by the population number in all grids in each CBG/city. Then, the Gini index[60] of NTL per person in the CBG/city boundary is computed as the light Gini. The NTL is the Visible Infrared Imaging Radiometer Suite (VIIRS) data[29,54] with a spatial resolution of 15 arc seconds (500 m at the Equator). We download the VIIRS Nighttime Lights version 2 Median monthly radiance (the unit for light intensity is nW /cm2/sr) with background masked from EOG[29,61–63]. Compared with income Gini from traditional income survey data, light Gini measures the NTL inequality in urban regions by considering NTL as an indicator for economic development, which is a different measurement of inequality[54].

Indicators for SDG 11 "Sustainable cities and communities". SDG 11 aims to make cities and human settlements inclusive, safe, resilient, and sustainable[3]. We incorporate indicators related to the built environment and land use in the target dataset. Specifically, we generate building density, driving/cycling/walking road density, POI density, land use information, and residential segregation (index of dissimilarity and entropy index) as indicators to monitor SDG 11.

The source data of urban built environment and land use is collected from OSM[31,64,65]. We collect the U.S. state-level historical Protocolbuffer Binary Format files from Geofabrik[32] from 2014 to 2023. Then we apply Python packet *pyrosm* to extract the building, driving road, cycling road, walking road, POI, and land use information in cities and CBGs by corresponding boundary polygons. For calculating building density, we divide the number of buildings by the area of CBG/city. For each of the three kinds of road density, we divide the total length of each kind of road by the corresponding area of CBG/city. The POI density, which is defined as the ratio of the number of all POIs and the area of CBG/city, can show urban venues with human information. The OSM POIs include all OSM elements with tags "amenity", "shop" or "tourism". The amenity tag is useful and important facilities for the urban population, which include Sustenance, Education, Transportation, Financial, Healthcare, Entertainment, Arts & Culture, Public Services, Waste Management, and Others. The shop tag includes locations of all kinds of shops and the sold products, such as Food & Beverages, General Store, Mall, Clothing, Shoes, Accessories, Furniture, etc. The tourism tag is the places for tourists, such as Museum, Gallery, Theme Park, Zoo, etc. Moreover, we generate the land use indicators (commercial, industrial, construction, and residential) by calculating the area percentage of each kind of land use in the area of CBG/city.

The indicators for the built environment quantitatively measure the density of buildings and roads. It should be noted that the indicators for SDG 11 are imperfect since the actual quality of buildings and roads is not provided in the dataset. Users can use the building/road/POI indicators as side information for depicting urban development.

Residential segregation is related to inclusivity in U.S. cities[66]. We calculate the index of dissimilarity[67]

$$D = \frac{1}{2} \sum_{i=1}^{n} \left| \frac{w_i}{w_T} - \frac{b_i}{b_T} \right|, \tag{1}$$

where n is the number of CBGs in a city, w_i is the number of race "w" (e.g., White) in CBG i, w_T is the total number of race "w" in the city, b_i is the number of race "b" (e.g., Black) in CBG i, and B_T is the total number of race "b" in the city. We calculate the index of dissimilarity for four racial or ethnic groups: Non-Hispanic White (White), Non-Hispanic Black or African American (Black), Non-Hispanic Asian (Asian), and Hispanic[66]. There are altogether six categories of indices of dissimilarity: White-Black, White-Asian, White-Hispanic, Black-Asian, Black-Hispanic, and Asian-Hispanic.

Next, we calculate the entropy index[68]

$$h_i = - \sum_{j=1}^{k} p_{ij} ln\left(p_{ij}\right), \tag{2}$$

where k is the number of racial/ethnic groups, p_{ij} is the proportion of jth race/ethnicity in CBG/city i. We include groups of the White, Black, Asian, and Hispanic population at the CBG or city level.

Limitations. The limitations of our dataset include errors from multiple data sources, partial coverage of SDG progress, and the shortcomings of selected indicators.

The errors from data sources include measurement errors in satellite imagery, ACS data collection, OSM, WorldPop population, and NTL data. The measurement errors in satellite imagery processing are mainly from the object detection and semantic segmentation tasks, and the accuracy metrics are shown in Table 5. And the errors in other data sources are usually tolerable in each field and the quality assessment can be referred to literature[69] for ACS data, literature[70–73] for OSM, literature[74] for WorldPop, and literature[62] for NTL data. ACS data uses sampling error to measure the difference between the true values for the entire population and the estimate based on the sample population. And the magnitude of sampling error is measured by the margin of error[69]. ACS provides a margin of error for all ACS estimate data which we collect as SDG indicators in our dataset. The dataset users can freely access the margin of error values of the ACS-oriented indicators in our dataset from the

Model	Dataset	Accuracy(%)	Precision(%)	Recall(%)	mAP@0.5(%)	mIoU(%)
Object Detection	xView	66.7	53.2	37.7	37.1	—
	DOTA v2	47.7	77.0	51.8	58.2	—
Semantic Segmentation	LoveDA	71.1	—	—	—	52.7

Table 5. Evaluation metrics for the object detection and semantic segmentation models.

ACS official website. OSM data is a Volunteered Geographic Information (VGI) and is frequently updated by volunteers. In terms of the road network, OSM is about 83% complete globally[70]. The building completeness for OSM in San Jose city in the U.S. is about 72% and confirms the validity of OSM building density in our dataset[71]. Some cities show a large jump in building number in a consecutive year due to lagging annotations. The POIs in OSM are compared with the Foursquare POIs and 60% of the POIs can be matched with high accuracy[72]. At last, the accuracy of the OSM land use dataset[73] for the U.S. is above 60%. The population data from WorldPop has a coefficients of determination[75] R^2 greater than 0.95 when evaluated on the population data in China[74]. The nighttime light intensity also shows a high consistency (R^2 greater than 0.97) compared with different nighttime light datasets[62].

And the provided dataset does not cover the whole SDG aspects, and thus cannot be used as the sole measurement for SDG monitoring. However, the dataset still has great reference value and aids decision-making for urban researchers and policymakers.

At last, some indicators cannot always be the best indicators for corresponding SDGs. For example, the indicator health insurance for SDG 3 (Good health and well-being) may not be the best measurement of health status because health insurance usage is affected by the income or wealth of the insurance owners.

Data Records

The produced dataset can be accessed through the Figshare repository[76] and is stored in tabular format. The live version with potential updates is available in the GitHub repository (https://github.com/axin1301/Satellite-imagery-dataset). We split the output dataset into seven categories, as shown in Fig. 2: basic geographic statistics, satellite imagery attributes, and five SDGs described in Fig. 1 to help users quickly access and utilize the data. Moreover, for each category, the dataset also contains records at two spatial levels (city and CBG). First, to help the users understand the area-of-interests, we provide samples of the geographic lookup table between the cities and CBGs in Table 2, and the basic statistics of CBGs and cities in Tables 6, 7. Second, to demonstrate the extracted visual attributes from the satellite imagery, we show samples of objects detected and land cover semantics from the satellite imagery at the CBG level in Table 8. At last, the samples of SDG indicators at the CBG level are demonstrated in Tables 9–13, respectively, which include SDG 1 "No poverty", SDG 3 "Good health and well-being", SDG 4 "Quality education", SDG 10 "Reduced inequalities", and SDG 11 "Sustainable cities and communities". The city name and CBG code are used to mark the geographical location of each SDG indicator.

Data table formats. *Basic geographical statistics.* Tables 6, 7 provide the population, area, geographic centroid, and geographic coordinate boundary of the area-of-interests in this dataset, where the area, geographic centroid, and boundary are invariant to time, while the population is time-varying.

Satellite imagery attributes. We have object numbers and land cover semantic attributes processed from satellite imagery of the years 2014 to 2023. The object categories include planes, airports, passenger vehicles, trucks, railway vehicles, ships, engineering vehicles, bridges, roundabouts, vehicle lots, swimming pools, soccer fields, basketball courts, ground track fields, baseball diamonds, tennis courts, and buildings (number of buildings). The land cover semantic attributes contain background, building (pixel percentage), road, water, barren, forest, and agriculture. There are altogether 24 visual attributes obtained from satellite imagery, which are shown in Table 8, where the objects detected consist of 17 columns and the land cover semantics consist of 7 columns. For visualization convenience, we only show the samples at the CBG level.

SDG 1. We provide median household income, population above/below poverty, population with a ratio of income to poverty level under 0.5, and population with a ratio of income to poverty level between 0.5 to 0.99 for "No poverty" indicators in Table 9 for the years 2014 to 2023 at the CBG level.

SDG 3. We offer civilian noninstitutionalized population, population with no health insurance under 18, population with no health insurance between 18 to 34, population with no health insurance between 35 to 64, and population with no health insurance over 65 years old for "Good health and well-being" indicators in Table 10 for the years 2014 to 2023 at the CBG level.

SDG 4. We provide population enrolled in college, population that graduated from high school, population with a bachelor's degree, a master's degree, and a doctorate for "Quality education" indicators in Table 11 for the years 2014 to 2023 at the CBG level.

SDG 10. We provide income Gini and light Gini for "Reduced inequalities" indicators in Table 12 for the years 2014 to 2023 at the CBG level. The income Gini measures the regional inequality from the perspective of

Geographic Area			Basic Geographical Statistics			
City Name	CBG Code	Year	Population	Area	Centroid	Boundary
Albuquerque	350010001071	2014	1,965	2.50	(−106.48718, 35.12321)	(−106.49789 35.13066, −106.49730 35.13067,…)
Albuquerque	350010001071	2015	1,992	2.50	(−106.48718, 35.12321)	(−106.49789 35.13066, −106.49730 35.13067,…)
Albuquerque	350010001071	2016	2,020	2.50	(−106.48718, 35.12321)	(−106.49789 35.13066, −106.49730 35.13067,…)
…						

Table 6. Example of basic geographic statistics of CBGs in the produced dataset.

Geographic Area			Basic Geographical Statistics			
City Name	City GeoID	Year	Population	Area	Centroid	Boundary
Albuquerque	1600000US3502000	2014	577,889	489	(−106.64648, 35.10534)	(−106.64882 35.14807, −106.64878 35.14818,…)
Albuquerque	1600000US3502000	2015	585,825	489	(−106.64648, 35.10534)	(−106.64882 35.14807, −106.64878 35.14818,…)
Albuquerque	1600000US3502000	2016	593,571	489	(−106.64648, 35.10534)	(−106.64882 35.14807, −106.64878 35.14818,…)
…						

Table 7. Example of basic geographical statistics of cities in the produced dataset.

Geographic Area			Objects Detected			Land Cover Semantics		
City Name	CBG Code	Year	Passenger Vehicle	Swimming Pool	…	Road	Forest	…
Albuquerque	350010001071	2014	411	48	…	0.0391	0.395	…
Albuquerque	350010001072	2014	111	42	…	0.0712	0.00979	…
Albuquerque	350010001081	2014	273	28	…	0.124	0.000537	…
…								

Table 8. Example of visual attributes from satellite images at the CBG level in the produced dataset.

Geographic Area			SDG 1 (No Poverty)				
City Name	CBG Code	Year	Median Household Income	Population Above Poverty	Population Below Poverty	Population With A Ratio Of Income To Poverty Level Under 0.5	Population With A Ratio Of Income To Poverty Level 0.5 to 0.99
Albuquerque	350010001071	2014	70,625	1,358	169	61	108
Albuquerque	350010001071	2015	79,276	1,293	133	57	76
Albuquerque	350010001071	2016	95,000	1,531	151	81	70
…							

Table 9. Example of indicators for SDG 1 at the CBG level in the produced dataset.

income, and the light Gini shows the regional nighttime light inequality through remote sensing technology. Since the income Gini at the CBG level is not available in ACS data, the corresponding values are omitted.

SDG 11. We provide building density, driving/cycling/walking road density, POI density, land use information (commercial, industrial, construction, and residential), and residential segregation (index of dissimilarity and entropy index) for "Sustainable cities and communities" indicators in Table 13 for the years 2014 to 2023 at the CBG level, where the index of dissimilarity has 6 columns: White-Black/ White-Asian/ White-Hispanic/ Black-Asian/ Black-Hispanic/ Asian-Hispanic. Since there is no data for proportions of the population with different races or ethnicities at the CBG level, the index of dissimilarity for CBG is omitted.

Technical Validation

Population percentage of city-of-interests. Our dataset selects the 100 most populated cities in the contiguous United States. We demonstrate the comparison of overall population in our city-of-interests and in all U.S. cities[33] in Fig. 3. We find that the population in selected cities takes up 52% of the total population in U.S. cities.

Visual attributes extraction from satellite imagery. We use state-of-the-art object detection and semantic segmentation models in the computer vision community to extract the visual attributes from satellite imagery. The training datasets, i.e., xView, DOTA v2, and LoveDA datasets, for the deep learning models are frequently used in satellite imagery interpretation tasks. We prepare the training datasets according to the models' requirements and transfer the trained models to the satellite images we collect. Following the evaluation methods

Geographic Area			SDG 3 (Good Health and Well-being)				
City Name	CBG Code	Year	Civilian Noninstitutionalized Population	Population With No Health Insurance Under 18	Population With No Health Insurance Between 18 To 34	Population With No Health Insurance Between 35 To 64	Population With No Health Insurance Over 65 Years Old
Albuquerque	350010001071	2014	1,527	0	10	62	0
Albuquerque	350010001071	2015	1,426	0	8	67	0
Albuquerque	350010001071	2016	1,682	0	22	38	0
...							

Table 10. Example of indicators for SDG 3 at the CBG level in the produced dataset.

Geographic Area			SDG 4 (Quality Education)				
City Name	CBG Code	Year	Population Enrolled In College	Population That Graduated From High School	Population With A Bachelor's Degree	Population With A Master's Degree	Population With A Doctorate
Albuquerque	350010001071	2014	77	171	267	242	108
Albuquerque	350010001071	2015	50	140	233	259	97
Albuquerque	350010001071	2016	77	207	276	250	108
...							

Table 11. Example of indicators for SDG 4 at the CBG level in the produced dataset.

Geographic Area			SDG 10 (Reduced Inequalities)	
City Name	CBG Code	Year	Light Gini	Income Gini
Albuquerque	350010001071	2014	0.761	—
Albuquerque	350010001071	2015	0.405	—
Albuquerque	350010001071	2016	0.392	—
...				

Table 12. Example of indicators for SDG 10 at the CBG level in the produced dataset.

Geographic Area			SDG 11 (Sustainable Cities and Communities)					
City Name	CBG Code	Year	Building Density	Driving/Cycling/ Walking Road Density	POI Density	Land Use	Index of Dissimilarity	Entropy Index
Albuquerque	350010001071	2014	20.3	8.28/9.63/9.87	1.59	0/0/0/33.7%	—	0.702
Albuquerque	350010001071	2015	39.9	8.87/11.6/11.9	1.99	0/0/0/33.7%	—	0.674
Albuquerque	350010001071	2016	39.9	8.71/11.5/11.7	1.99	0/0/0/33.7%	—	0.786
...								

Table 13. Example of indicators for SDG 11 at the CBG level in the produced dataset.

in the computer vision community[27,77,78], we also present the evaluation metrics for the object detection and semantic segmentation models on the evaluation datasets in Table 5. Specifically, we show the accuracy, precision, and recall for all object categories, as well as the mean Average Precision under Intersection over Union (IoU) threshold 0.5 (mAP@0.5) for the object detection models on xView and DOTA v2 datasets and the accuracy and mean IoU (mIoU) for the semantic segmentation model on LoveDA dataset. For purposes of comparison, the state-of-the-art model SAC[79] reaches a mAP@0.5 of 27.2%, which is lower than the 37.1% of our object detection model on xView dataset. For the DOTA v2.0 dataset, the best mAP is obtained by DCFL[80] reaching 57.66%. However, since the image split and experiment settings in the object detection model training processes might be different, the direct comparison of performance metrics of our models and the state-of-the-art literature on xView and DOTA v2.0 datasets is for reference only. For the semantic segmentation task, the mIoU of our model 52.8% is close to the state-of-the-art model UperNet[81] with an mIoU of 52.44%. Such results guarantee the usefulness and credibility of our produced data. In addition, to test the robustness of our trained models qualitatively, we randomly select satellite images with their corresponding object detection and semantic segmentation results, which are shown in Fig. 4. We visualize the object detection results in satellite imagery in Fig. 4a, where buildings and passenger vehicles are identified. For the satellite imagery semantic segmentation model, the ViT-Adapter-based model shows high performance in recent semantic segmentation tasks, and the example of segmentation result is shown in Fig. 4b. The results prove the effectiveness of transferring the pre-trained models to our satellite imagery.

SDG indicators prediction from satellite imagery visual attributes. Visual information in satellite imagery correlate with income/daily consumption[11,14,82], commercial activity[16,83–86], education level[12], and
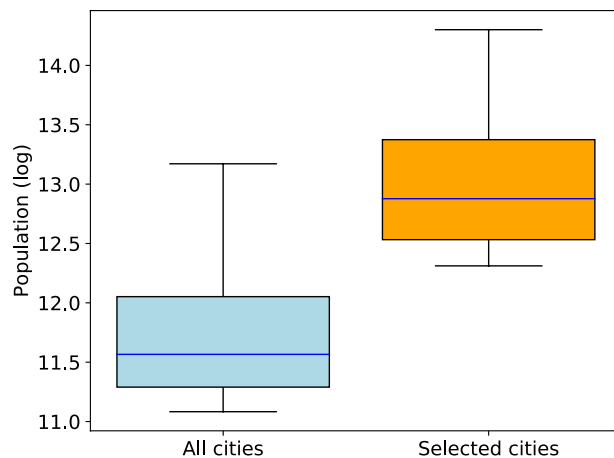
Fig. 3 Comparison of the overall population in city-of-interests in our dataset and population in all U.S. cities.
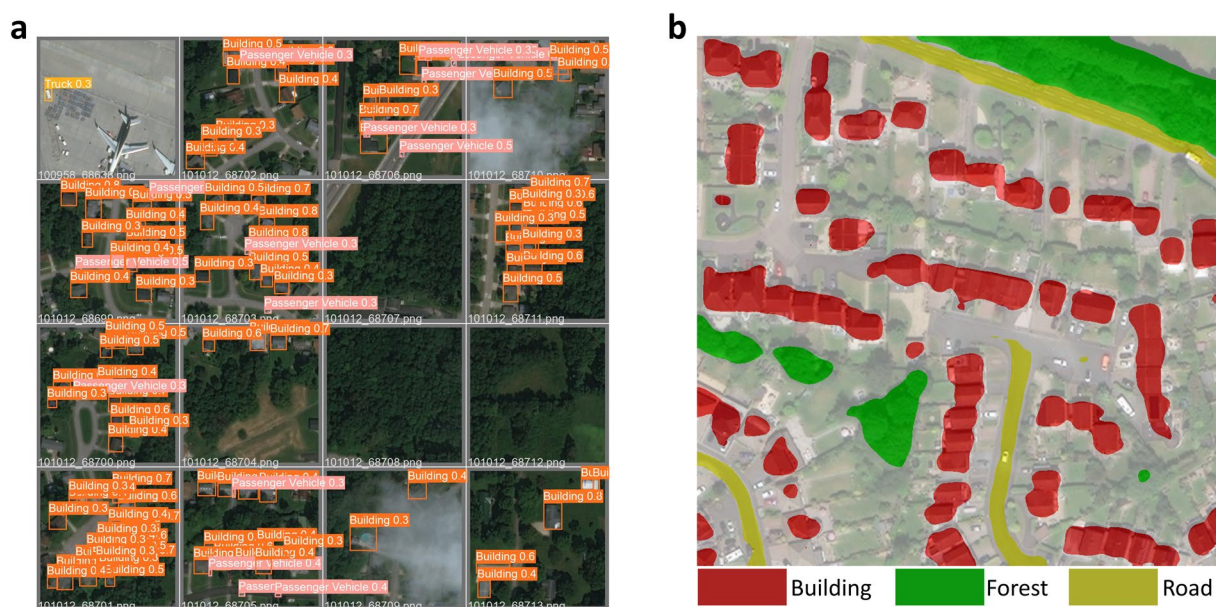

Fig. 4 Visualization of the (**a**) object detection and (**b**) semantic segmentation results. (Zoom-in is recommended to visualize the bounding box classes in (**a**)).

health outcome[12,87]. Therefore, we validate the possibility of inferring SDG indicators from corresponding satellite imagery. Specifically, for each CBG, the visual attributes (see Table 3) of satellite imagery are fed into a regression model to infer the SDG indicators. We select median household income, population with no health insurance at all ages, population that graduated from high school, and POI density in 2018 as the indicators for poverty, health status, education, and commercial activity, respectively. We experiment on whether those indicators can be inferred from the satellite images by applying Gradient Boosting Decision Trees (GBDT)[88] on the satellite imagery visual attributes and the indicators selected above as output at the CBG level. The ground truth data for training and validation of the regression models are the collected SDG indicators in our dataset. We randomly split the 100 cities into 80 training cities and 20 validation cities, and thus all the CBGs in one city are grouped into the same fold. The regression results are shown in Fig. 5, where we can see that the coefficient of determination[75] R^2 of the predicted median household income and POI density with regard to the ground truth are higher than the R^2 for health and education indicators. Specifically, GBDT has a prediction performance of R^2 reaching about 0.155 and 0.338 for median household income and POI density, respectively. These results are consistent with the findings in previous research[11,14,16,82] that socioeconomic status can be inferred from satellite imagery, confirming the validity of the provided dataset and demonstrating the potential to monitor SDGs from satellite imagery. While the education and health indicators are predicted with low precision, which encourages dataset users for future enhancement. At present, most research on predicting socioeconomic status from satellite imagery focuses on income/poverty (SDG 1) and commercial activity (POI density in SDG 11). The studies for inferring regional health or education status are very few, and the performance of relevant prediction models is much lower than the
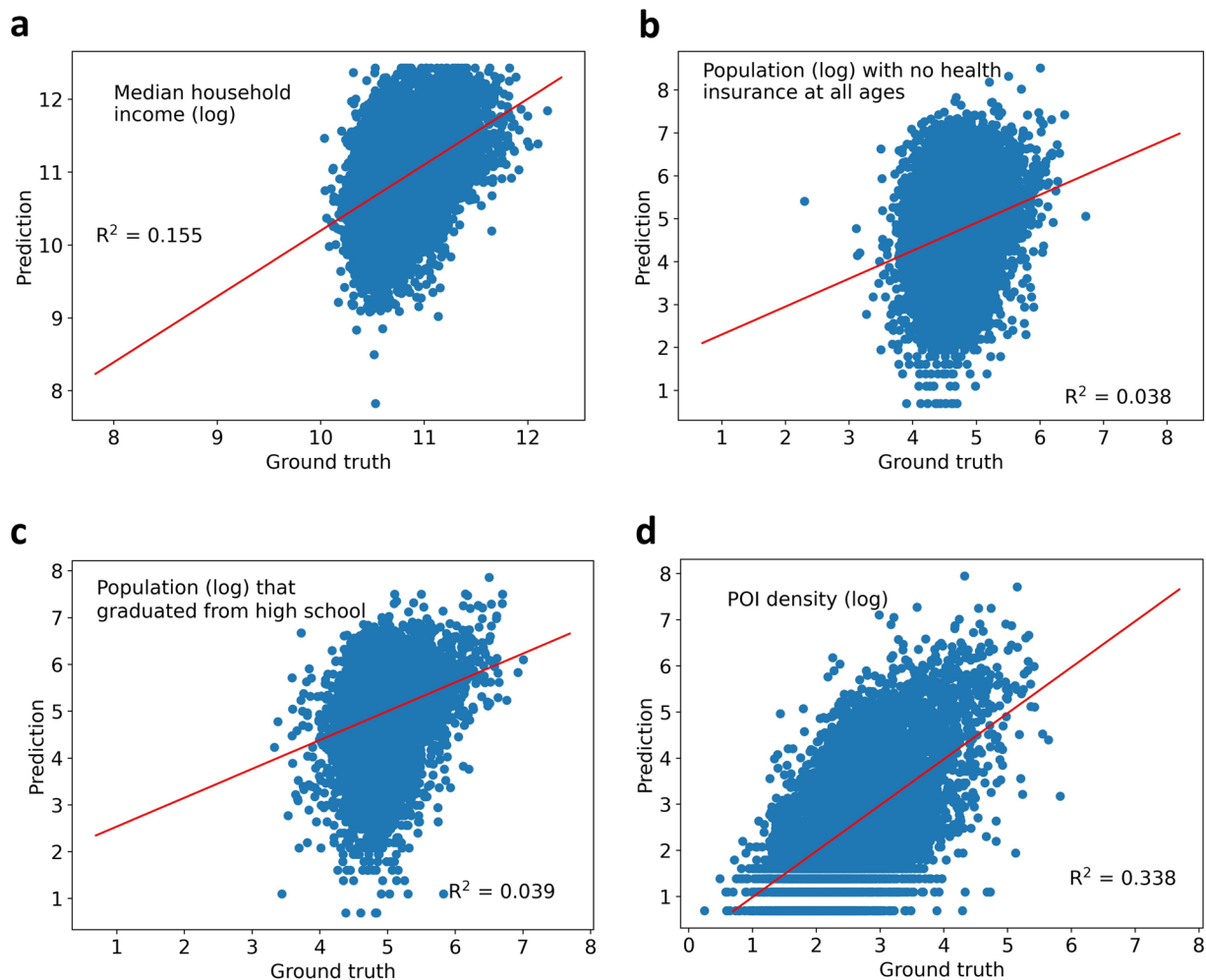
Fig. 5 SDG indicators prediction results from satellite imagery visual attributes with GBDT at the CBG level in 2018. (**a**) Median household income (log), (**b**) Population (log) with no health insurance at all ages, (**c**) Population (log) that graduated from high school, and (**d**) POI density (log).

performance of income prediction (see Fig. 5), which makes the health and education-related SDG monitoring a promising research direction in the future.

Usage Notes

This study aims to provide a long-term and multi-scale dataset in cities covering the satellite imagery attributes and SDG indicators for urban policymakers and researchers to advance SDG monitoring. Specifically, the satellite imagery attributes in our dataset can be used as input for proposing machine learning models to predict the SDG indicators. Moreover, the SDG data in our dataset can also provide insights into how SDGs evolve in time or scale. Since our dataset contains various aspects of cities, we recommend the following potential research applications: introducing new methods for predicting poverty/income, health, education, inequality, and living environment status of people in cities from long-term or multi-scale satellite images. Researchers are also encouraged to discover the underlying relationship between various SDG progresses and satellite images in cities.

The dataset files at the CBG level has about 400,000 lines of data, which might take a long time to load in Excel. Thus, we recommend loading the data with a Python script that can handle large datasets. For the object numbers inferred from the satellite imagery, the confidence level above 0.2 is counted.

Code availability

The Python codes to collect, process, and plot the dataset as well as the supplementary files for this study are publicly available through the GitHub repository (https://github.com/axin1301/Satellite-imagery-dataset). Detailed instruction for the running environment, file structure, and codes is available in the repository.

References

1. The World Bank. Urban development. https://www.worldbank.org/en/topic/urbandevelopment/overview (2022).
2. United Nations Department of Economic and Social Affairs. World urbanization prospects: The 2018 revision. https://www.un-ilibrary.org/content/books/9789210043144 (2019).
3. United Nations. Transforming our world: The 2030 agenda for sustainable development. https://sdgs.un.org/publications/transforming-our-world-2030-agenda-sustainable-development-17981 (2015).
4. Alainna Lynch and Jeffrey Sachs. The United States Sustainable Development Report 2021. https://us-states.sdgindex.org/ (2021).
5. Prakash, M. The U.S. Cities Sustainable Development Goals Index 2017. https://www.sdgindex.org/reports/2017-u.s.-cities-sdg-index/ (2017).
6. United States Census Bureau. American Community Survey data. https://www.census.gov/programs-surveys/acs (2022).
7. Safegraph. Open-Census-Data. https://www.safegraph.com/free-data/open-census-data (2022).
8. United States Census Bureau. Fiscal Year 2022 Budget Summary U.S. Census Bureau. https://www2.census.gov/about/budget/census-fy-22-budget-infographic-bureau-overview.pdf (2022).
9. United Nations Sustainable Development Solutions Network. "data for development: A needs assessment for SDG monitoring and statistical capacity development". https://resources.unsdsn.org/data-for-development-a-needs-assessment-for-sdg-monitoring-and-statistical-capacity-development (2015).
10. Burke, M., Driscoll, A., Lobell, D. B. & Ermon, S. Using satellite imagery to understand and promote sustainable development. *Science* **371**, eabe8628 (2021).
11. Jean, N. *et al.* Combining satellite imagery and machine learning to predict poverty. *Science* **353**, 790–794 (2016).
12. Head, A., Manguin, M., Tran, N. & Blumenstock, J. E. Can human development be measured with satellite imagery? In *Ictd*, 8–1 (2017).
13. Chen, C. *et al.* Analysis of regional economic development based on land use and land cover change information derived from landsat imagery. *Scientific Reports* **10**, 1–16 (2020).
14. Ayush, K., Uzkent, B., Burke, M., Lobell, D. & Ermon, S. Generating interpretable poverty maps using object detection in satellite images. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, IJCAI'20 (2021).
15. Albert, A., Kaur, J. & Gonzalez, M. C. Using convolutional networks and satellite imagery to identify patterns in urban environments at a large scale. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1357–1366 (2017).
16. Wang, W. *et al.* Urban perception of commercial activeness from satellite images and streetscapes. In *Companion Proceedings of the The Web Conference 2018*, 647–654 (2018).
17. Yeh, C. *et al.* Using publicly available satellite imagery and deep learning to understand economic well-being in africa. *Nature communications* **11**, 1–11 (2020).
18. Sumbul, G., Charfuelan, M., Demir, B. & Markl, V. Bigearthnet: A large-scale benchmark archive for remote sensing image understanding. In *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*, 5901–5904 (IEEE, 2019).
19. Van Etten, A., Lindenbaum, D. & Bacastow, T. M. Spacenet: A remote sensing dataset and challenge series. Preprint at https://arxiv.org/abs/1807.01232 (2018).
20. Irvin, J. *et al.* Forestnet: Classifying drivers of deforestation in indonesia using deep learning on satellite imagery. In *Thirty-fourth Conference on Neural Information Processing Systems Workshop on Tackling Climate Change with Machine Learning* (2020).
21. The United Nations Educational, Scientific and Cultural Organization (UNESCO). Unesco launches 2022 survey on public access to information. https://www.unesco.org/en/articles/unesco-launches-2022-survey-public-access-information (2022).
22. Yeh, C. *et al.* SustainBench: Benchmarks for Monitoring the Sustainable Development Goals with Machine Learning. In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)* (2021).
23. Lam, D. *et al.* xView: Objects in context in overhead imagery. Preprint at https://arxiv.org/abs/1802.07856 (2018).
24. Ding, J., Xue, N., Long, Y., Xia, G.-S. & Lu, Q. Learning roi transformer for oriented object detection in aerial images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2849–2858 (2019).
25. Xia, G.-S. *et al.* Dota: A large-scale dataset for object detection in aerial images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 3974–3983 (2018).
26. Ding, J. *et al.* Object detection in aerial images: A large-scale benchmark and challenges. *IEEE transactions on pattern analysis and machine intelligence* **44**, 7778–7796 (2021).
27. Wang, J., Zheng, Z., Ma, A., Lu, X. & Zhong, Y. Loveda: A remote sensing land-cover dataset for domain adaptive semantic segmentation. In Vanschoren, J. & Yeung, S. (eds.) *Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks*, vol. 1 (Curran, 2021).
28. Guenat, S. *et al.* Meeting sustainable development goals via robotics and autonomous systems. *Nature Communications* **13**, 1–10 (2022).
29. Elvidge, C. D., Baugh, K., Zhizhin, M., Hsu, F. C. & Ghosh, T. VIIRS night-time lights. *International Journal of Remote Sensing* **38**, 5860–5879 (2017).
30. Tatem, A. J. Worldpop, open data for spatial demography. *Scientific Data* **4**, 1–4 (2017).
31. OpenStreetMap Foundation & Contributors. Openstreetmap. https://www.openstreetmap.org/ (2022).
32. Geofabrik GmbH, OpenStreetMap Foundation & Contributors. Geofabrik downloads. http://download.geofabrik.de/north-america/us.html (2022).
33. United States Census Bureau. American Community Survey table s0101: Sex and age. https://data.census.gov/table?q=S0101&g=0100000US$1600000y=2021 (2021).
34. United States Census Bureau. 2020 TIGER/Line Shapefiles. https://www2.census.gov/geo/tiger/TIGER2020/PLACE/ (2020).
35. He, Z., Yang, S., Zhang, W. & Zhang, J. Perceiving commerial activeness over satellite images. In *Companion Proceedings of the The Web Conference 2018*, WWW '18, 387–394 (2018).
36. Han, S. *et al.* Learning to score economic development from satellite imagery. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2970–2979 (2020).
37. Esri. World Imagery. https://livingatlas.arcgis.com/wayback/ (2022).
38. Knowles, R. D., Ferbrache, F. & Nikitas, A. Transport's historical, contemporary and future role in shaping urban development: Re-evaluating transit oriented development. *Cities* **99**, 102607 (2020).
39. Sustainable Development Goals Knowledge Platform. Sustainable transport. https://sustainabledevelopment.un.org/topics/sustainabletransport (2022).
40. Carlucci, M., Vinci, S., Ricciardo Lamonica, G. & Salvati, L. Socio-spatial disparities and the crisis: Swimming pools as a proxy of class segregation in athens. *Social Indicators Research* 1–25 (2020).
41. Wang, H., Dai, X., Wu, J., Wu, X. & Nie, X. Influence of urban green open space on residents' physical activity in china. *BMC public health* **19**, 1–12 (2019).
42. Wilfried Lemke. The role of sport in achieving the sustainable development goals. https://www.un.org/en/chronicle/article/role-sport-achieving-sustainable-development-goals (2016).
43. Abitbol, J. L. & Karsai, M. Interpretable socioeconomic status inference from aerial imagery through urban patterns. *Nature Machine Intelligence* **2**, 684–692 (2020).

44. Chen, L. *et al*. Uvlens: urban village boundary identification and population estimation leveraging open government data. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* **5**, 1–26 (2021).

45. Jocher, G. *et al*. ultralytics/yolov5: v7.0 - YOLOv5 SOTA Realtime Instance Segmentation. *Zenodo* https://doi.org/10.5281/zenodo.7347926 (2022).

46. Jocher, G. *et al*. The official implementation of the yolov5 model. https://github.com/ultralytics/yolov5 (2022).

47. Lin, T.-Y. *et al*. Microsoft coco: Common objects in context. In *Computer Vision–ECCV 2014*: *13th European Conference*, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13, 740–755 (Springer, 2014).

48. Jocher, G. *et al*. Hyperparameters of the yolov5 model. https://github.com/ultralytics/yolov5/blob/master/train.py (2022).

49. Chen, Z. *et al*. Vision Transformer Adapter for Dense Predictions. In *The Eleventh International Conference on Learning Representations* (2023).

50. Chen, Zhe. The official implementation of "Vision Transformer Adapter for Dense Predictions". https://github.com/czczup/ViT-Adapter/tree/main/segmentation (2022).

51. Chen, Z. *et al*. Hyperparameters of "Vision Transformer Adapter for Dense Predictions". https://github.com/czczup/ViT-Adapter/blob/main/segmentation/configs/ade20k/upernet_augreg_adapter_base_512_160k_ade20k.py (2022).

52. Zhou, B. *et al*. Scene Parsing through ADE20K dataset. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, 633–641 (2017).

53. Worldpop Hub. The spatial distribution of population with country total adjusted to match the corresponding unpd estimate, united states. https://hub.worldpop.org/geodata/listing?id=69 (2022).

54. Mirza, M. U., Xu, C., van Bavel, B., van Nes, E. H. & Scheffer, M. Global inequality remotely sensed. *Proceedings of the National Academy of Sciences* **118** (2021).

55. United States Census Bureau. How the census bureau measures poverty. https://www.census.gov/topics/income-poverty/poverty/guidance/poverty-measures.html (2023).

56. United States Census Bureau. American Ccommunity Survey Information Guide. https://www.census.gov/content/dam/Census/programs-surveys/acs/about/ACS_Information_Guide.pdf (2017).

57. United States Census Bureau. American Community Survey. https://data.census.gov/all?q=B19013 (2022).

58. Pan, J., Lei, X. & Liu, G. G. Health insurance and health status: exploring the causal effect from a policy intervention. *Health economics* **25**, 1389–1402 (2016).

59. Meng, Y., Han, J. & Qin, S. The impact of health insurance policy on the health of the senior floating population–evidence from china. *International journal of environmental research and public health* **15**, 2159 (2018).

60. Damgaard, C. & Weiner, J. Describing inequality in plant size or fecundity. *Ecology* **81**, 1139–1142 (2000).

61. Elvidge, C. D., Baugh, K. E., Zhizhin, M. & Hsu, F.-C. Why VIIRS data are superior to DMSP for mapping nighttime lights. *Proceedings of the Asia-Pacific Advanced Network* **35**, 62 (2013).

62. Elvidge, C. D., Zhizhin, M., Ghosh, T., Hsu, F.-C. & Taneja, J. Annual time series of global VIIRS nighttime lights derived from monthly averages: 2012 to 2019. *Remote Sensing* **13**, 922 (2021).

63. Earth Observation Group. VIIRS nighttime lights (VNL) version 2 median monthly radiance with background masked. https://eogdata.mines.edu/nighttime_light/annual/v21/ (2022).

64. Haklay, M. & Weber, P. Openstreetmap: User-generated street maps. *IEEE Pervasive computing* **7**, 12–18 (2008).

65. Vargas-Munoz, J. E., Srivastava, S., Tuia, D. & Falcao, A. X. Openstreetmap: Challenges and opportunities in machine learning and remote sensing. *IEEE Geoscience and Remote Sensing Magazine* **9**, 184–199 (2020).

66. Chodrow, P. S. Structure and information in spatial segregation. *Proceedings of the National Academy of Sciences* **114**, 11591–11596 (2017).

67. Sakoda, J. M. A generalized index of dissimilarity. *Demography* **18**, 245–250 (1981).

68. Iceland, J. The multigroup entropy index (also known as theil's h or the information theory index). *US Census Bureau. Retrieved July 31*, 60 (2004).

69. United States Census Bureau. 7. understanding error and determining statistical significance. https://www.census.gov/content/dam/Census/library/publications/2018/acs/acs_general_handbook_2018_ch07.pdf (2018).

70. Barrington-Leigh, C. & Millard-Ball, A. The world's user-generated road map is more than 80% complete. *PloS one* **12**, e0180698 (2017).

71. Zhou, Q. Exploring the relationship between density and completeness of urban building data in openstreetmap for quality estimation. *International Journal of Geographical Information Science* **32**, 257–281 (2018).

72. Zhang, L. & Pfoser, D. Using openstreetmap point-of-interest data to model urban change–a feasibility study. *PloS one* **14**, e0212606 (2019).

73. Zhou, Q., Wang, S. & Liu, Y. Exploring the accuracy and completeness patterns of global land-cover/land-use data in openstreetmap. *Applied Geography* **145**, 102742 (2022).

74. Ma, J. *et al*. Accuracy assessment of two global gridded population dataset: A case study in china. In *Proceedings of the 4th International Conference on Information Science and Systems*, 120–125 (2021).

75. Lewis-Beck, C. & Lewis-Beck, M. *Applied regression: An introduction*, **vol. 22** (Sage publications, 2015).

76. Xi, Y. *et al*. A satellite imagery dataset for long-term sustainable development in United States cities, *figshare*, https://doi.org/10.6084/m9.figshare.c.6425261.v1 (2023).

77. Tan, M., Pang, R. & Le, Q. V. Efficientdet: Scalable and efficient object detection. In *Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition*, 10781–10790 (2020).

78. Padilla, R., Netto, S. L. & da Silva, E. A. B. A survey on performance metrics for object-detection algorithms. In *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, 237–242 (2020).

79. Liu, J., Levine, A., Lau, C. P., Chellappa, R. & Feizi, S. Segment and complete: Defending object detectors against adversarial patch attacks with robust patch detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 14973–14982 (2022).

80. Xu, C. *et al*. Dynamic coarse-to-fine learning for oriented tiny object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 7318–7328 (2023).

81. Wang, D. *et al*. Advancing plain vision transformer towards remote sensing foundation model. *IEEE TGRS* (2022).

82. Han, S. *et al*. Lightweight and robust representation of economic scales from satellite imagery. *Proceedings of the AAAI Conference on Artificial Intelligence* **34**, 428–436 (2020).

83. Xi, Y. *et al*. Beyond the first law of geography: Learning representations of satellite imagery by leveraging point-of-interests. In *Proceedings of the ACM Web Conference 2022*, 3308–3316 (2022).

84. Li, T. *et al*. Predicting multi-level socioeconomic indicators from structural urban imagery. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, 3282–3291 (2022).

85. Li, T. *et al*. Learning representations of satellite imagery by leveraging point-of-interests. *ACM Transactions on Intelligent Systems and Technology* **14**, 1–32 (2023).

86. Liu, Y., Zhang, X., Ding, J., Xi, Y. & Li, Y. Knowledge-infused contrastive learning for urban imagery-based socioeconomic prediction. In *Proceedings of the ACM Web Conference 2023*, 4150–4160 (2023).

87. Levy, J. J. *et al*. Using satellite images and deep learning to identify associations between county-level mortality and residential neighborhood features proximal to schools: A cross-sectional study. *Frontiers in Public Health* **9**, 766707 (2021).

88. Friedman, J. H. Greedy function approximation: a gradient boosting machine. *Annals of statistics* 1189–1232 (2001).

Author contributions
Yong Li, Tong Li, Yu Liu, Yanxin Xi and Pan Hui contributed to conceptualizing the study. Yanxin Xi acquired raw data, produced the dataset, and plotted the figures. Yunke Zhang collected the data partly. Yu Liu and Yanxin Xi contributed to the drafting of the manuscript. All authors revised the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to Yong Li or P.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.