AI model trained with images can recognize visual indicators of gentrification

March 5 2024, by Katharine Miller

Credit: Pixabay/CC0 Public Domain

Seemingly overnight, a yoga studio replaces a barbershop, a coffee café takes over a small grocery, and a multi-story apartment building looms where older single-family homes once stood.

Such signs of gentrification can be jarring to longtime city residents. They know that by the time these physical upgrades have happened, it could be too late for local policymakers to protect them from soaring rents and shifts in who can afford to live in the neighborhood.

Yet despite being highly visible, gentrification and its consequences have been difficult to measure or study at scale, says Jackelyn Hwang, assistant professor of sociology at Stanford University's School of Humanities and Sciences.

Researchers have had to rely on demographic data from either the U.S. Census, captured once a decade, or the American Community Survey (ACS), which happens every five years and isn't very precise. In addition, the data from these sources is aggregated by census tract rather than building by building. "It's not sufficiently fine-grained," Hwang says.

Now, Hwang, a Stanford HAI affiliate, and graduate student Tianyuan Huang describe training a model to recognize the visual cues of gentrification—typically new construction or renovations—in Google Street View images for entire cities.

In recent work the team showed that about 74% of the time the model predicted gentrification in the same places where gentrification had been previously found in studies of census and ACS demographic data. In addition, the model found early signals of gentrification in neighboring census blocks previously labeled as non-gentrifying.

"Detecting change at the image level really complements the demographic changes we've seen in other data that's not as granular," Hwang says.

In the long run, more granularity around where gentrification is happening will allow cities to target anti-displacement remedies more precisely, Hwang says. For example, better information could prompt cities to pass rent protections or just-cause eviction rules, target affordable housing development or rent subsidies, or place caps on property tax increases for low-income long-term residents.

"Being able to measure patterns of reinvestment in places that are gentrifying as well as disinvestment in places that don't gentrify can help us think about allocating resources in ways that would better help vulnerable communities," Hwang says.

The gentrification research challenge

Gentrification, defined as the socio-economic upgrading of previously low-income urban spaces, is poorly understood. For example, though population displacement is considered one of the most significant consequences of gentrification, quantitative research finds that gentrification and displacement are unrelated, Hwang says.

The challenge is this: If gentrification is causing displacement, then neighborhoods that don't gentrify would experience less displacement, which isn't the case, she says. "What we see is that most evictions and other forced moves are taking place in the non-gentrifying low-income neighborhoods that don't receive investment."

But high mobility rates among low-income, predominantly Black neighborhoods that don't gentrify might still be a result of upward rent pressures caused by gentrification nearby, Hwang says. "Prices go up everywhere to a different extent, so even if higher income people aren't moving into these neighborhoods, landlords may still be reacting to price increases across the city." To get a better handle on these dynamics, there's a need for more granular data.

The gentrification detection model

To address that problem, Hwang and Huang and their colleagues turned to Google Street View, a data set of street-level images covering all of the major and minor cities in the United States and across the world. First created in 2007, Google Street View is updated every one to three years. Moreover, the data is not aggregated: It includes images of each stretch of road, each building, and each empty lot.

For their training dataset, the researchers pulled construction permits to identify where construction was planned, and extracted data on business upscaling (laundry to coffee shop; grocery to high-end restaurant) from a national business directory. After confirming these positive examples of gentrification by manually looking at paired before/after images in the Google Street View image dataset, they trained a model to recognize building upgrades.

The model used a "self-attention mechanism" to make sure the gentrification signal wouldn't get lost in the crowd. "Like a human, when shown a bunch of images it tries to focus on the meaningful ones," Huang says. In this case, the attention layer looks first for buildings and then looks closer to see if there were improvements to the building from one point in time to the next. It then assigns greater weight to the images that draw its attention.

One challenge of the work was defining an appropriate way to validate the model. There is no ground truth measure of gentrification (indeed, that's the problem here), but the team decided to compare the model to prior attempts to label census tracts as gentrifying or non-gentrifying using demographic data from the U.S. Census and the ACS.

In a pilot study using pairs of images (from 2007 through 2022) from the full Google Street View data set for three cities (Oakland, Denver, and

Seattle), the model labeled the same gentrifying or non-gentrifying census tracts as the prior work 74% of the time.

Interestingly, the model identified a significant number of what could be false positives—census tracts the model labeled as gentrifying that had been labeled non-gentrifying in the past. These could have been errors by the model, but when the researchers looked at paired images in those census tracts, they found what looked like gentrification—new apartment buildings and neighborhood upgrades.

The conclusion: Because the model leveraged granular street-level imagery, it seemed to be spotting early signs of gentrification that previous studies had missed.

The team's gentrification model was published in *IEEE Big Data* in late 2022, but a year later the team simplified the model to be a predictor of construction (rather than gentrification) and used time-series images rather than a single pair of images to help the system distinguish between temporary changes (seasons, shadows, camera angle) and real ones.

As a construction predictor, the model was 90% accurate. Moreover, the model's construction predictions were highly correlated with both population growth and income growth, suggesting a clear tie to gentrification.

"If there are new apartments, the neighborhood population grows very fast, and that's a change that can be quite visible," Huang says. The datasets are public and the new work is posted on the *arXiv* preprint server, but it has been accepted for publication by the Association for the Advancement of Artificial Intelligence (AAAI).

Together, the two papers suggest a promising future for the use of Google Street View images to understand urban change.

Next, Hwang is working on another Google Street View project to study how changes to the built environment impact individuals' health and community well-being.

Meanwhile, Huang is using Google Street View to spot not only building improvements but also building demolition and decline in communities after extreme climate events like hurricanes, floods, and wildfires. "The goal," Huang says, "is to see if those places are being rebuilt or if people have abandoned them."

By having larger-scale data, the researchers can look at variation across different disasters in different places rather than just one disaster site at a time. "This should help us to better understand which communities are benefiting from reinvestment and which are not, as well as who is benefiting and who is not," Hwang says.

More information: Tianyuan Huang et al, CityPulse: Fine-Grained Assessment of Urban Change with Street View Time Series, *arXiv* (2024). DOI: 10.48550/arxiv.2401.01107

Tianyuan Huang et al, Detecting Neighborhood Gentrification at Scale via Street-level Visual Data, *2022 IEEE International Conference on Big Data (Big Data)* (2023). DOI: 10.1109/BigData55660.2022.10020341

Provided by Stanford University