Using Deep Learning and Google Street View to Estimate the Demographic Makeup of the US

Timnit Gebru, Jonathan Krause, Yilun Wang, Duyun Chen, Jia Deng, Erez Lieberman Aiden, Li Fei-Fei

Presented by
Chia-Wen Cheng
Wed Nov 8, 2017
Each year, the U.S. Census Bureau spends $1 billion surveying the population.
Challenges of Population Survey

- Labor-intensive
- Time-consuming
- Ignore smaller areas
A faster, more efficient, and higher-resolution way of studying the population?
The type of car people own is a strong indicator of their demographic information.
Vehicular Census via Google Street View Images

- 200 American cities
- 50 million Street View Images
- 22 million vehicles
- 2,657 different categories of cars
- Vehicle characteristics
  - Make, model, year, body type...
Automated System for Monitoring Demographic Trends

- Street View images
- Car Detection
- Car Classification
- Demographic Estimation
- Extract car-related attributes

race, income, education, voting pattern
Car Detection

- Deformable Part Models (DPMs)
- Tradeoff between performance and efficiency

Image credit: P. Felzenszwalb et al.
Car Classification

Street View images

Product shot images

AlexNet

2657-way classification

domain adaptation
Car-Related Attributes

88 attributes:

- The average number of detected cars per image
- Average car price
- Miles per gallon
- Percent of total cars that are hybrids
- Percent of total cars that are electric
- Percent of total cars that are from each of seven countries
- Percent of total cars that are foreign (not from the USA)
- Percent of total cars from each of 11 body types
- Percent of total cars whose make is each of 58 makes in our dataset
Demographic Estimation

88 car-related attributes

Ridge regression

Logistic regression

Income
Voter preference

Race
Education
Results

Race

i. White (Seattle, Washington)
ii. Black (Seattle, Washington)
iii. Asian (Seattle, Washington)
Results

Education

Income
An interesting finding

“ If the number of sedans encountered during a 15-minute drive through a city is higher than the number of pickup trucks, the city is likely to vote for a Democrat during the next Presidential election; otherwise, it is likely to vote Republican. “
Strengths

Overall:

● Solve a practical problem
● Think outside of the box

Technique:

● Almost real-time monitoring
● Fine spatial resolution
Weaknesses

- Hand-crafted car-related attributes
- Correlation between car ownership and demographics?
- Generalizable to other demographic factors?
  e.g. religion, birth rate, death rate, marriage age, marital status
Extension

- Other types of imagery
  - e.g. Drone View images, satellite images
- Predict traffic flow