A Deep Learning Approach for Population Estimation from Satellite Imagery

Caleb Robinson, Fred Hohman, Bistra Dilkina

ACM SIGSPATIAL Workshop on Geospatial Humanities
Redondo Beach, California, United States
Knowing Where People Live is Important
Knowing Where People Live is Important
Knowing Where People Live is Important

- Urban development
- Disaster response
Knowing Where People Live is Important

- Urban development
- Disaster response
- Infectious disease containment
Sustainable Development Goals

Most rely on population counts for measuring outcomes

1. **No Poverty**
   - Need to know population of a country

2. **Good Health and Well-being**
   - Need to measure population mortality rates

3. **Reduced Inequalities**
   - Need to know where people live in relation to services
Geospatial Humanities Applications

What is the influence of physical or geographical space on human behavior and cultural development? [1]

Using detailed population data and other accessory pieces of data we can learn relationships that may be useful in historic or cultural contexts.

Relationship between:

- **land-cover** and **population** with the Historic Land Dynamics Assessment
- **population distribution in cities** - **cultural differences in cities**

Censuses are Hard!

Expensive, time consuming, varying degrees of accuracy, varying degrees of spatial resolution, ...

2010 round - 5 countries without a census

2000 round - 27 countries without census

Lebanon hasn't taken a census since 1932
Much can Change Between Censuses!

Time-lapse of Las Vegas, Nevada

Made with: https://earthengine.google.com/timelapse/
How Satellite Data Can Help

Models based on satellite data are:

- *Much faster* than waiting on census
- Can be used in-between census years
- Can be applied to other countries without population counts
- Highly scalable

Landsat
The Landsat project (NASA and USGS) has been continuously collecting satellite images of the entire earth since 1972.

*Landsat 7 has been in operation since 1999, taking images of the entire globe every 16 days.*
The Landsat project (NASA and USGS) has been continuously collecting satellite images of the entire earth since 1972.

*Landsat 7* has been in operation since 1999, taking images of the entire globe every 16 days.

What can we learn from all of this data?
- What is in a single satellite image?
- Where people live
Landsat 7 imagery of Redondo Beach (2010)
We are learning to estimate population from satellite imagery.

Satellite Image to Population Mapping

Convolutional Neural Network

100 people
Gridded Population Data Products

Gridded Population of the World
Global Rural-Urban Mapping Project
WorldPop
Facebook Labs

LandScan

Population Ground Truth → Disaggregated to grid
Gridded Population Data Products

Gridded Population of the World  LandScan
Global Rural-Urban Mapping Project  WorldPop  Facebook Labs

Need to know ground truth!
Gridded Population Estimations

Equitable development through deep learning:
The case of sub-national population density estimation

Patrick Doupe
Arnold Institute for Global Health
New York, NY, USA
patrick.doupe@msn.com

James Fahmy
Arnold Institute for Global Health
New York, NY, USA
james.fahmy@msn.com

Emilie Bruzelius
Arnold Institute for Global Health
New York, NY, USA
emilie.bruzelius@msn.com

Samuel G. Ruchman
Office of the UN Secretary-General’s Special Envoy for Health in Agenda 2030 and for Malaria
New York, NY, USA
sruchman@ omnipower.org

ABSTRACT

High-resolution population density maps are a critical component for global development efforts, including service delivery, resource allocation, and disaster response. Traditional population density efforts are predominantly survey driven, which are laborious, prohibitively expensive, infrequently updated, and inaccurate—especially in remote areas. Furthermore, these maps are developed on a regional basis where the methods used vary region to region, hence introducing notable spatio-temporal heterogeneity and bias.

The advent of global-scale satellite imagery provides us with an unparalleled opportunity to create inexpensive, accurate, homogeneous, and rapidly updated population maps. To fully realize this vision, we must overcome both infrastructure and methodological obstacles. We propose a convolutional neural network approach that addresses some of the methodological challenges, while employing a publicly available, global low resolution, remote sensed product. The method converts satellite imagery into population density estimates. To explore both the accuracy and generalizability of our approach, we train our neural network on Tanzanian imagery and test the model on Kenyan data. We show that our method is able to generalize to unseen data and we improve upon the current state of the art by 17 percent.

1. INTRODUCTION

The 2015 adoption of the Sustainable Development Goals has created an acute need to collect accurate subnational population statistics to inform global development plans and policies. Furthermore, subnational data are a critical component of equitable development since national averages routinely "smooth out" the most marginalized communities. In fact, one of the most significant limitations to achieving the Millennium Development Goals was that, while national averages have shown overall gains towards certain goals, these gains veered significant disparities, where the poorest people did not necessarily benefit from the overall progress [36]. One of the most basic, yet essential, of these statistics is the accurate and timely assessment of where people live. This information is "one of the primary sources of data needed for formulating, implementing and monitoring the effectiveness of policies and programmes aimed at inclusive socioeconomic development and environmental sustainability" [40, p.2]. These census data determine resource allocation, such as where to invest in hospitals, schools and infrastructure, and may be used to define legislative districts and other important functional units of government. Their population level data also serve as an essential benchmark for measuring progress toward the attainment of the Sustainable Development Goals, and other national and international objectives (see Figure 1).

Traditional population density estimates are derived from census surveys. However, the utility of censuses is hampered by several widely recognized quality issues. First, they are "among the most complex and massive enterprise exercises a nation undertakes" [42, p.5]. Many countries still lack the capacity, both in terms of financial and human resources, to collect data regularly [7]. As a result, population estimates are largely based on "extrapolated or inferred demographic facts" [36, p.4].

Gridded Population Estimations

Equitable development through deep learning:
The case of sub-national population density estimation

Patrick Doupe
Arnold Institute for Global Health
New York, NY, USA
p_doupe@amnh.edu

James Faghmous
Arnold Institute for Global Health
New York, NY, USA
jfaghmous@amnh.edu

Emilie Bruzelius
Arnold Institute for Global Health
New York, NY, USA
emilie.bruzelius@amnh.edu

Samuel G. Ruchman
Office of the UN Secretary-General’s Special Envoy for Health in Agenda 2030 and for Malaria
New York, NY, USA
sruchman@unesco.org

ABSTRACT
High-resolution population density maps are a critical component for global development efforts. Using service delivery, resource allocation, and disaster response. Traditional population density efforts are predominantly survey driven, which are laborious, prohibitively expensive, infrequently updated, and inaccurate – especially in remote areas. Furthermore, these maps are developed on a regional basis where the methods used vary from region to region, hence introducing notable equi-temperatures heterogeneity and bias. The advent of global-scale satellite imagery provides us with an unprecedented opportunity to create inexpensive, accurate, homogenous, and rapidly updated population maps. To fulfill this vision, we must overcome both infrastructure and methodological obstacles. We propose a convolutional neural network approach that addresses some of the methodological challenges, while employing a publicly available, all-lower-resolution, remote sensed product. The method converts satellite imagery into population density estimates. To explore both the accuracy and provability of our approach, we train our neural network on Tanzanian imagery and test the model on Kenya data. We show that our method is able to generalize to unseen data and we improve upon the current state of the art by 17 percent.

1. INTRODUCTION
The 2035 adoption of the Sustainable Development Goals has created an acute need to collect accurate sub-national population density estimates. Furthermore, subnational data are a critical component of equitable development since national averages routinely “smooth out” the most marginalized communities. In fact, one of the most significant limitations to achieving the Millennium Development Goals was that, while national averages may have showed overall gains towards certain goals, these gains were not necessarily significant disparities, where the poorest people did not necessarily benefit from the overall progress [26]. One of the most basic, yet essential, of these statistics is the accurate and timely assessment of where people live. This information is “one of the primary sources of data needed for formulating, implementing, and monitoring the effectiveness of policies and programs aimed at inclusive socioeconomic development and environmental sustainability” [40, p.2]. These census data determine resource allocation, such as where to invest in hospitals, schools, and infrastructure, and may be used to define legislative districts and other important functional areas of government. Their population level data also serve as an essential benchmark for measuring progress toward the attainment of the Sustainable Development Goals, and other national and international objectives (see Figure 1). Traditional population density estimates are derived from census surveys. However, the utility of censuses is hampered by several widely recognized quality issues. First, they are “among the most complex and massive pieces of data a nation undertakes” [42, p.5]. Many countries still lack the capacity, both in terms of financial and human resources, to collect data regularly [7]. As a result, populations are left to suffer the consequences of being undercounted, or overcounted.
Remaining Outline

- Deep learning overview
- Data
- Methods
- Results
- Future work
Deep learning is a category of state-of-the-art algorithms that has recently been shown to be very effective for image-based learning.

Example: convolutional neural networks (CNNs)

Goal: learn a mapping between satellite images and population counts

Data

*Landsat 7* images

- $74 \times 74 \times 7$ "images" representing $1\text{km}^2$ patches of land

Gridded US Census population count data

- Disaggregate block group data onto grid
- Raw population counts
- Population ranges

For every $1\text{km}^2$ patch of land in the US we have:

- a satellite image, existing population estimate for years 2000 and 2010
Methods

Estimating population ranges instead of counts (classification v. regression)

Class 0: 0 people, Classes 1-7: few people, Classes 8-13: many people

Train on data from 2000, test on data from 2010

How do we validate our model?

• Aggregate gridded estimations at county level - compare to ground truth
• Show that it makes reasonable predictions
Methods

Sample training and validation sets from the 2000 data → Training → Evaluate on validation set → Evaluate on 2010 test data

Architecture/Parameter Search

Pixel level error

County level error
Neural Network Architecture

Input Images
Each image of size (74, 74, 7)

Modified VGG-A Neural Network

Output
Classification vector of size (17)
Quantitative Results

Results are at the county level

- Our models: **CONVRAW** and **CONVAUG**
- Census Bureau estimates: **POSTCENSAL** and **ACS5YR**

<table>
<thead>
<tr>
<th>Method</th>
<th>Mean Absolute Error</th>
<th>Median Absolute Error</th>
<th>$r^2$</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONVRAW</td>
<td>23,005</td>
<td>6,357</td>
<td>0.91</td>
<td>73.78</td>
</tr>
<tr>
<td>CONVAUG</td>
<td>19,484</td>
<td>4,642</td>
<td>0.94</td>
<td>49.82</td>
</tr>
<tr>
<td>POSTCENSAL</td>
<td>2,020</td>
<td>559</td>
<td>0.99</td>
<td>3.09</td>
</tr>
<tr>
<td>ACS5YR</td>
<td>1,704</td>
<td>214</td>
<td>0.99</td>
<td>34.44</td>
</tr>
</tbody>
</table>

2010 Mean Population: 97,018
2010 Median Population: 25,930

Difference between Actual 2010 and CONVAUG 2010
Qualitative Interpretation

1. Which satellite images are the models most confident about?

2. Where do the models predict low → high population areas are?

3. What errors are the models making?
1. Which satellite images are the models most confident about?

Top 8 most confident test images for each population class.
2. Where do the models predict low → high population areas are?
3. What errors are the models making?

- **Oak Ridge National Lab**
  - *Oak Ridge, TN*
  - Small Scale

- **Anniston Army Depot**
  - *Anniston, AL*
  - Medium Scale

- **Walt Disney World**
  - *Orlando, FL*
  - Large Scale
Future Work

Can we improve the accuracy of the models?
- More recent neural network architectures
- Effects of imperfect training data
- Loss function improvements

Do our models generalize in space?
- Test a model trained on the US in different countries?

Can we predict other socio-economic values?
A Deep Learning Approach for Population Estimation from Satellite Imagery

Thanks!

Fred Hohman
fredhohman.com
fredhohman@gatech.edu

Caleb Robinson
calebrob.com
dcrobins@gatech.edu

Bistra Dilkina
cc.gatech.edu/~bdilkina
bdilkina@cc.gatech.edu

We thank NSF for funding this work and the anonymous reviewers for their constructive feedback.