Estimating GDP

And Economic Growth

From Satellite Imaging

Team Members:

Robert Lopez-perez

Austin Semple

Utkarsha Anil Gurjar

Vaidehi Sonar

Abhijeet Pokhriyal

2.1.1 Business/Customers

- What problem will your Computer Vision solution solve, and for whom?

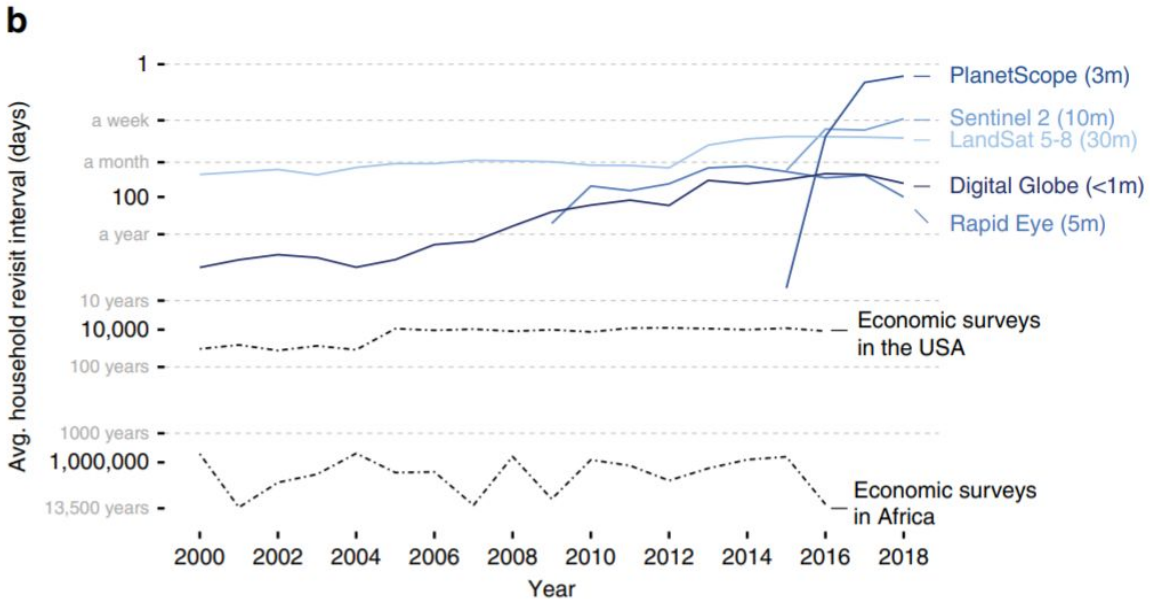
  Consider an unlikely problem: finding the poor. Even in a world riddled with poverty, nearly every government, nonprofit and aid agency struggles with this issue.
  Proxy for finding the poor can be basic economic variables like gross domestic product. But the problem is that these numbers can be unreliable in countries where the statistical infrastructure is weak, the informal businesses do not want to be tracked and the numbers may be manipulated.[DEATON & HESTON]
  It is very difficult to randomly sample people in the rural areas of Bihar in India or in a slum like Kibera in Nairobi, Kenya, where even just mapping the streets is its own project.[NYT]
  In most countries GDP numbers are not available on any consistent basis at the subnational level. Much of the interesting variation in economic growth takes place within, rather than between, countries and that is what can help us with our original problem of poverty.[HENDERSON]
  Computer vision can help by using Satellite images with Nighttime luminosity. This tells us not just about electrification but also about economic activity more broadly, and statistical work shows it reliably correlates with economic performance. [HENDERSON]

- What value will it provide them? What are their pain points?

  Value provided by using Satellite data instead of extensive surveying is that we can use publicly available imagery to infer both spatial and temporal differences in local-level economic well-being specially for countries where reliable survey data do not yet exist and where survey-based interpolation methods might struggle to generate accurate estimates.

  Pain points that this approach can address are

  a. Expense - Land surveys are sampled surveys and involve extensive investments of time and money. On the other hand satellites can provide more up-to-date information and a way faster pace
  b. Limited repeated observation - these surveys have limited repeated observation of individual locations, making it difficult to measure local changes in well-being over time, and public release of any disaggregated consumption data from is very rare.[Yeh, C., Perez]



- How big is the potential market?

With this approach we can get a handle on the level of economic activity at any level international, national or subnational/ district level. This information can be of great help not just to organizations like IMF, WHO but also to local businesses looking to expand and scale.

The potential market is as big as the number of (developing) countries. The approach is easily scalable and limited only by the processing power of computers. With the advent of big data processing platforms like Hadoop, we can go from raw satellite data to valuable insights in no time.

2.1.2 Academic Literature Review

A thorough literature review can save weeks of wasted time implementing dead-end solutions or re-doing work that others have already done. You should pick 3 academic papers (we can help you with this, if needed), and do a "deep-dive" on these publications.

- What academic work is relevant to your project topic? Pick 3 papers, ask us for help if you need it.

1. Sungwon Han, Donghyun Ahn, Sungwon Park, Jeasurk Yang, Susang Lee, Jihee Kim, Hyunjoo Yang, Sangyoon Park, and Meeyoung Cha. 2020. Learning to Score Economic Development from Satellite Imagery.

- https://dl.acm.org/doi/10.1145/3394486.3403347

2. Yeh, C., Perez, A., Driscoll, A. *et al.* Using publicly available satellite imagery and deep learning to understand economic well-being in Africa. *Nat Commun* 11, 2583 (2020).
    - https://doi.org/10.1038/s41467-020-16185-w
3. Henderson, J. Vernon, Adam Storeygard, and David N. Weil. 2012. "Measuring Economic Growth from Outer Space." *American Economic Review*, 102 (2): 994-1028.
    - https://www.aeaweb.org/articles?id=10.1257/aer.102.2.994

- What makes these papers important/relevant?

  These papers are important because they not just outline the general methodology of how to proceed with the use of satellite images for predicting economic activity but also offer varying degrees of feature engineering approaches that can be used. All of them use publicly available data and offer comparative study between the developing and the developed nations.

- What are their results and how did they achieve these results?

  Many of their results exceed performance in earlier work on a similar task using high-resolution imagery, or mobile phone data as input, and match or exceed benchmarks for in-country performance from geostatistical models used to predict health outcomes, standard of living, and housing quality. They are able to achieve these results using a mix of statistical and deep learning models. They incorporate both sets of imagery (daytime and night light) in a deep learning model trained end-to-end, with models trained separately on daytime and nighttime images and then joined in a final fully connected layer. The goal of the model is to learn features in the daytime and nighttime imagery that are predictive of asset wealth, without first prescribing what features the model should look for.

- What's different/unique about these approaches?

  One of the most distinguishing features of one of the papers is that it's semi-supervised and therefore does not presuppose availability of labelled image data. One of the papers tries to not just predict but understand why some locations are wealthier than others. They study associations between wealth and exposure to extreme temperatures, they also use the estimates to evaluate the hypothetical targeting of a social protection program (e.g. a cash transfer).

2.1.3 Open Source

- What open source code is available that is relevant to your topic?

1. Earthpy- https://earthpy.readthedocs.io/en/latest/gallery_vignettes/index.html
    a. EarthPy is a python package devoted to working with spatial and remote sensing data. EarthPy also contains an IO module that supports downloading data for the Earth Lab earth analytics courses and any user with a url and a zip file.
2. Rasterio - https://rasterio.readthedocs.io/en/latest/quickstart.html
    a. Geographic information systems use GeoTIFF and other formats to organize and store gridded raster datasets such as satellite imagery and terrain models. Rasterio reads and writes these formats and provides a Python API based on Numpy N-dimensional arrays and GeoJSON.
3. Raster vision - https://github.com/azavea/raster-vision
    a. This open source project makes it easy for teams to build computer vision models to understand and analyze geospatial imagery.
4. Google Earth Engine - https://code.earthengine.google.com/
    a. Not open source but allows easy/ free access to many of the datasets using a convenient API

- How active are the communities around this code?
    - Earthpy is fairly well maintained with a small community of active developers and contributors.
    - Rasterio is a mapbox product and therefore more widely used with around 100 contributors
    - Azavea - https://www.azavea.com/work/
        - Azavea offers a suite of open source tools and services that help in the development of geospatial analytical tools

- What data is available for testing and/or training algorithms?

1. Bhuvan App - https://bhuvan-app1.nrsc.gov.in/
    a. It is an Indian web based utility which allows users to explore a set of map based content prepared by Indian Space Research Organisation. The content which the utility serves is mostly restricted to Indian boundaries
2. DMSP-OLS Nighttime Lights
    a. https://eogdata.mines.edu/dmsp/downloadV4composites.html
    b. These files are cloud-free composites made using all the available archived DMSP-OLS smooth resolution data for calendar years 1992-2013
3. VIIRS Day/Night Band Nighttime Lights

    a. https://eogdata.mines.edu/download_dnb_composites.html
    b. The VIIRS Nighttime Imagery (**Day/Night Band**, Enhanced Near Constant Contrast) layer shows **the Earth's** surface and atmosphere using a sensor designed to capture low-light emission sources, under varying illumination conditions. It is displayed as a grey-scale image.
4. International Monetary Fund
    a. https://www.imf.org/external/datamapper/datasets
    b. IMF DataMapper publishes a wide selection of the key economic indicators from 11Datasets including GDP, Public Debt, World Economic outlook indicator etc

- Is labeled data available? How much? How is the data licensed? Is it under copyright protection?

  The image dataset available are unlabeled geospatial images. But to produce a labelled dataset we plan to include IMF datasets for the corresponding regions and timeframes. The datasets are not licensed and can be used freely.

2.1.4 Industry Solutions

- What companies are solving similar problems to yours?
  A joint collaboration between Stamen Design, DigitalGlobe, and Carnegie Mellon University, is working on Penny which is artificial intelligence that can "read" satellite imagery of two very different cities and judge the income brackets of neighborhoods within them.

- If you have access to the product, what can you learn from using it?

  The product Penny is an experimental product. Penny is an implementation of **RESNET-50**, a deep convolutional neural network, trained on DigitalGlobe high-resolution satellite imagery and census block data. It is very interesting to know the ability of Penny to manipulate satellite imagery to try and raise or lower the predicted wealth in an area. You can move around tall buildings, add or remove trees and land and so on and can see how it affects Penny's predictions about the median income of this hypothetical new neighborhood.

- Are there available talks, documentation, or other resources from their engineering teams?

  There are a few articles or blogs available :

- https://www.bloomberg.com/news/articles/2017-06-30/artificial-intelligence-predicts-wealth-from-satellite-images
- https://www.digitaltrends.com/home/penny-machine-learning-income-predictor/
- https://www.weforum.org/agenda/2017/10/penny-artificial-intelligence-satellite-designs-cities/

SOURCES

1. [Yeh, C., Perez]https://www.nature.com/articles/s41467-020-16185-w
    a. Yeh, C., Perez, A., Driscoll, A. *et al.* Using publicly available satellite imagery and deep learning to understand economic well-being in Africa. *Nat Commun* 11, 2583 (2020). https://doi.org/10.1038/s41467-020-16185-w
2. [NYT](https://www.nytimes.com/2016/04/03/upshot/satellite-images-can-pinpoint-poverty-where-surveys-cant.html)
3. [HENDERSON](https://www.aeaweb.org/articles?id=10.1257/aer.102.2.994)
4. [DEATON & HESTON](https://pubs-aeaweb-org.librarylink.uncc.edu/doi/pdfplus/10.1257%2Fmac.2.4.1)
5. [WORLD_BANK](https://blogs.worldbank.org/sustainablecities/tracking-light-space-innovative-ways-measure-economic-development)
6. [NASA](https://earthdata.nasa.gov/learn/sensing-our-planet/prosperity-shining)
7. [IMF](https://www.imf.org/external/pubs/ft/fandd/2019/09/satellite-images-at-night-and-economic-growth-yao.htm)