Using Satellite Imagery to Predict Multidimensional Poverty in Nigeria

Amy Jin*[1], Maximillian Guo*[1]

[1]Harvard University, Cambridge, MA
amy_jin@college.harvard.edu, mguo@college.harvard.edu

Abstract

In recent years, computer vision and satellite imagery have been used to predict poverty in various regions worldwide. However, existing works often utilize individual poverty indices such as income or consumption expenditure, as well as high-resolution, commercial satellite imagery. Our work presents these main contributions: 1) a self-curated dataset containing publicly accessible, lower resolution satellite image mosaics for each of the 37 regions of Nigeria, 2) an image sampling approach for data augmentation that requires imagery for only 0.01% of the country's land area, 3) a custom model that directly predicts poverty without relying on high-resolution images or intermediary proxies, and 4) prediction of the multidimensional poverty index (MPI), a more holistic measure of poverty that gives insight across multiple dimensions of poverty including health, education, and living standard indicators, in Nigeria. We find that a basic convolutional neural network is capable of accurately predicting MPI, especially when given additional structural information that constitutes MPI such as population in MPI and deprivation intensity or the ten components of MPI. This suggests that satellite imagery and sampling can be used to cheaply predict MPI and potentially other more holistic poverty measures across smaller regions or regions where it may be difficult to collect data.

Introduction

Nine percent of the world, close to 700 million people, live in extreme poverty on less than \$2.15 per day (World Bank 2022). Historically, poverty was predominantly measured by the proportion of a population living below the poverty line. However, in the 1990s, economists began to examine spatial dimensions to poverty, leading to the development of poverty maps, which enable richer understanding of where poverty is most acute, as well as where and how to intervene (Atia 2014) (Poverty and Initiative 2022b).

While individual poverty indices such as income or consumption expenditure are still commonly used today, the experience of poverty is more complex and can be felt across several dimensions, whether in health, education, or other factors. The multidimensional poverty index (MPI) was developed in 2010 by the Oxford Poverty and Human Development Initiative (OPHI) and United Nations Development Programme (UNDP). MPIs track deprivations across three dimensions and 10 indicators including health (nutrition, child mortality), education (years of schooling, school attendance), and living standards (cooking fuel, sanitation, drinking water, electricity, housing, assets), offering a higher resolution lens into where and how poverty affects people (Poverty and Initiative 2022a). Equation 1 delineates how each indicator contributes to the overall MPI of a region, where c_h, c_e, and c_l denote indicators within the health, education, and living standard dimensions, respectively. In looking beyond income, 1.3 billion people in 107 developing countries are multidimensionally poor (United Nations Development Programme 2020).

$$MPI = \frac{1}{6}(c_{h,1} + c_{h,2} + c_{e,1} + c_{e,2}) + \frac{1}{18}\left(\sum_{i=1}^{6} c_{l,i}\right) \quad (1)$$

Our motivations for undertaking this work are thus threefold. First, because existing studies such as (Ayush et al. 2021), (Blumenstock, Cadamuro, and On 2015), and (Yeh et al. 2020) commonly index on only a single poverty measure such as income or consumption expenditure, we leverage MPIs and poverty mapping to provide richer insight into poverty-stricken regions, better informing intervention programs and policies.

However, it can be difficult to collect such comprehensive data for each MPI indicator at granular, regional and subregional levels, as the data collection process typically involves consumption and census surveys, which are expensive in terms of costs, labor, and time. Thus, our work explores whether publicly accessible satellite imagery could be directly used to predict MPI.

Finally, we focus our work on Nigeria given its poverty profile, as 4 in 10 citizens live below the national poverty line, and many lack access to education and basic infrastructure (World Bank 2022).

Related Work

Previous works have leveraged satellite imagery or other remote sensing data for poverty prediction in various regions and types of environments. They largely differ along a few main axes, including prediction methodology, and poverty and satellite imagery data used.

*These authors contributed equally.

Poverty Prediction Methodologies

Several studies use proxies from satellite imagery or other remote sensing data to map poverty. For example, (Perez et al. 2017) and (Jean et al. 2016) employ nighttime light intensity as an intermediate proxy to predict poverty in five African countries. (Engstrom, Hersh, and Newhouse 2017) extract features from high-resolution satellite imagery such as number and density of buildings, number of cars, density and length of roads, and roof material to estimate poverty rates and consumption in Sri Lanka, and (Abelson, Varshney, and Sun 2014) similarly utilizes roof type as a proxy for poverty. (Babenko et al. 2017) explores the predictive power of land use estimates derived from satellite imagery.

Rather than relying on proxies, our work directly predicts poverty using satellite imagery, saving on intermediary computational steps and manual image annotations, which are computationally and labor intensive tasks.

Poverty Measures

To the best of our knowledge, this study is also the first to predict the multidimensional poverty index. Existing works more commonly use single-dimensional measures such as asset wealth index, household income, income per adult equivalent, and consumption expenditure (Jean et al. 2016) (Babenko et al. 2017) (Piaggesi et al. 2019) (Pandey, Agarwal, and Krishnan 2018). Our goal is to be able to give richer insight into different ways people can experience poverty across health, education, and living standards.

Satellite Imagery

Furthermore, previous works often rely on medium to high spatial resolution, frequently commercial, satellite imagery in order to accurately estimate poverty. Some commonly used sources include Planet and DigitalGlobe imagery, which provide commercial images at 3-5m and 50cm resolution, respectively, as well as Google Static Maps, which offers 2.5m spatial resolution (Babenko et al. 2017) (Piaggesi et al. 2019).

Our work leverages publicly available, lower resolution (10m) imagery from Sentinel-2 satellites and contributes a new dataset covering all of Nigeria. As Sentinel-2 satellites have imaged the globe continuously every five days since 2015, this data source provides the advantages of having a consistent, reliable collection and release timeline, in addition to reducing costs and computational complexity.

Dataset

We use two datasets for this work, one consisting of MPI data for Nigeria, and another which is a novel dataset we curate of publicly available Sentinel-2 satellite imagery spanning each of the 36 states of Nigeria and the Federal Capital Territory (FCT), for a total of 37 regions. We describe each one in further detail.

Multidimensional Poverty Indices

Our model is trained using multidimensional poverty index data for Nigeria for 2021, which is publicly released by OPHI and can be accessed through the Humanitarian Data Exchange (Humanitarian Data Exchange 2021). Along with the overall MPI for each region, it provides the contributions of each of the ten indicators that compose of the index, as well as population data and estimates of populations in poverty and deprivation intensity based on MPI. Figure 1 displays a choropleth map of overall MPI for Nigeria along with its distribution.

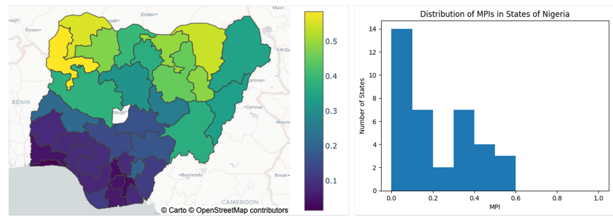

Figure 1: Choropleth map of overall MPI for Nigeria (left) along with its distribution (right).

Figure 2: A zoomed-out view of the cloudless Sentinel-2 satellite imagery obtained using Google Earth Engine for Nigeria.

Sentinel-2 Satellite Imagery

We also contribute a new dataset of satellite imagery for all of Nigeria. Using the Google Earth Engine API, we retrieve publicly accessible Sentinel-2 Level-1C tiles and create mosaics for each of the 37 regions (Google Earth Engine 2022). In order to obtain an unobstructed view of the land, we set the cloud coverage percentage to be less than one percent. Since image tiles meeting this cloud coverage criterion are often taken at different points in time, they vary in consistency for each visited location. To increase image consistency across tiles while constructing each region's mosaic, we create composite images for each region by sampling the median pixel value of each cloud-free tile for the years 2016 to 2022. Additionally, while the Sentinel-2 satellite imagery has 13 spectral bands in total, we combine only the B4, B3, and B2 bands, which correspond to red, green, and blue and have a spatial resolution of 10m. In total, our dataset consists of approximately 80 GB of image data. Figure 2 shows a high-level aggregate view of the 37 image mosaics, spanning all of Nigeria.

Methods

Given our limited dataset of only 37 regions, we use a method for augmenting our dataset that generalizes for any

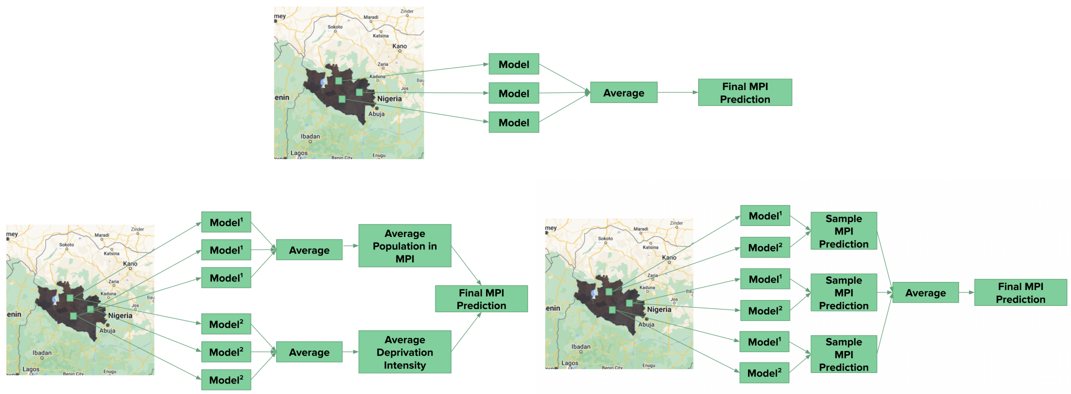

Figure 3: Different ways of obtaining MPI predictions. top: direct method, bottom left: factor method, average first, bottom right: factor method, average last. The components method is similar to the factor method but with different predictors.

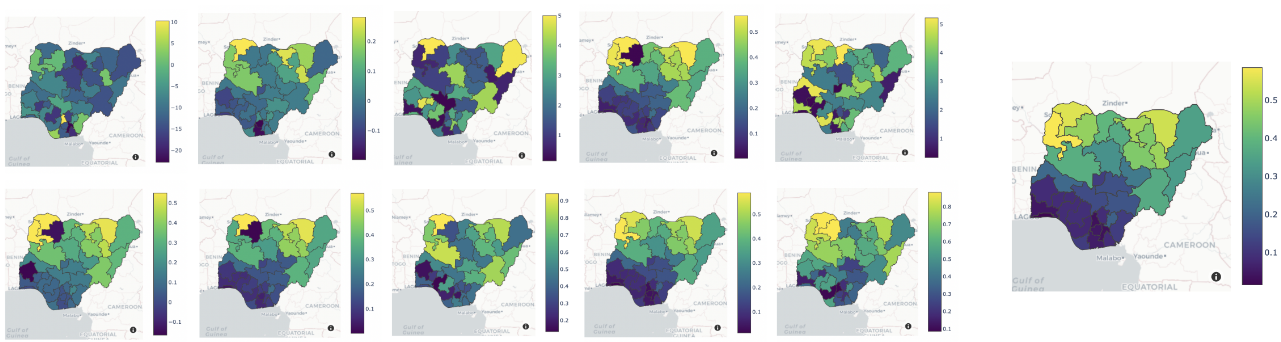

Figure 4: Choropleth map of true MPI (right most image) and the overall MPI predictions for each of the five methods (ten right images). From left to right: 1) direct method, 2) factor method, average first, 3) factor method, average last, 4) components, average first 5) components, average last. The top row is using linear regression, while the bottom row uses the CNN.

situation in which there is a plethora of input data but limited response data. For each region, we uniformly sample 12 regions of 40 by 40 pixels (corresponding to roughly 400m by 400m on the ground) of Sentinel-2 satellite imagery in that region and label each image in the same region with the same response value, where we experiment with the number of samples. This corresponds to roughly 0.01% of the country's land area. For each region, we use approximately 75% of the satellite imagery samples for training and the remaining 25% for testing. For our *direct method*, we directly predict the MPI as the label. Now to obtain a prediction for a state, we do an average of predictions of all MPI predictions for satellite imagery within the state. For our *factor method*, we predict separately both *population in MPI* and *deprivation intensity* and obtain MPI estimates via the following formula:

$$MPI = (population\ in\ MPI) \cdot (deprivation\ intensity) \quad (2)$$

However, note that we can either 1) compute for the entire state an average of population in MPI and an average of deprivation intensity and then use Equation (2) to obtain the MPI for the state or 2) compute MPI from population in MPI and deprivation intensity for each sample and then average the MPIs to obtain the MPI for the state. We denote these two sub-methods as *average first* (AF) and *average last* (AL), respectively. Lastly, similar to the *factor method*, we can perform the procedure separately for each of the components of MPI (nutrition, child mortality, ..., assets), then combine them using equation Equation (1), which we call the *components* method. This can also be done with an average first or average last procedure. An overview of our methodology is shown in Figure 3.

Models

For the predictive model that takes an input satellite imagery and outputs either MPI, a factor of MPI, or a component of MPI, we utilized a basic convolutional neural network with two convolutional layers and two max-pool layers. Full details are in the Appendix. As a baseline, we also implement a basic linear regression (LR) on the satellite pixel values.

Results and Discussion

	Test Loss for Samples (MSE)	
Response Value	Linear Regression	CNN
MPI	118.61	0.004
Population in MPI	641.75	127.87
Deprivation Intensity	809.48	0.088
Nutrition	444.79	1.973
Child Mortality	21.46	3.15
Years of Schooling	373.51	0.618
School Attendance	136.3	0.446
Cooking Fuel	242.21	0.992
Sanitation	225.18	0.084
Drinking Water	99.73	0.563
Electricity	153.83	0.023
Housing	362.78	0.359
Assets	69.04	0.835

Table 1: Results on samples from satellite imagery demonstrate that the CNN is much more capable than the baseline of predicting each response value. Note that some of the data are on different scales, so the MSE should be compared relatively between models.

	Test Loss for MPI (MSE)	
	Linear Regression	CNN
Direct	1.421	0.0126
Factor, AF	0.0557	0.00810
Factor, AL	6.492	0.0903
Components, AF	0.00656	0.000023
Components, AL	9.4393	0.0766

Table 2: Results for final MPI prediction for each of the 37 states using the various methods.

Given the varying contributions of the health, education, and living standards dimensions of poverty, we investigate how effectively our model can predict specific components of MPI in addition to overall MPI from satellite imagery. Table 1 shows the test results for MPI, population in MPI, deprivation intensity, and each of the ten components of MPI. Note that we did not scale the data, so MPI, Population in MPI, and Deprivation Intensity are each on their own scales (these quantities are sometimes measured in percentages, meaning that, if data is not scaled, all predictions must be scaled down by a factor of 100^2 when using formulas such as Equation (2). However, the ten components of MPI are necessarily on the same scale (Equation (1)). Of the 10 components of MPI, child mortality seems to be the most difficult to predict by the CNN, although interestingly enough it is the best response value to predict out of the ten components for linear regression. The linear regression still performs many orders of magnitude worse than the CNN, exhibiting the power of deep learning with satellite imagery.

Table 2 shows the results for our final MPI predictions using the different methods of direct prediction or combining predictions for factors or components. As a general trend, averaging first does better than averaging last across both models, suggesting that averaging first might be a method of mitigating extreme combinations of noisy predictions that might compound outlier error. Moreover, when using averaging first, the components method does better than the factor method, which does better than the direct method. In fact, the components method using averaging first and the CNN model matches the true MPI almost perfectly, suggesting that individual prediction errors are mitigated and smoothed out via Equation (1). However, this presents a trade-off in the real world. The results suggest that more components of data seem to lead to better MPI predictions (given structural equations Equation (1) or Equation (2)), but this requires more work to obtain.

These results demonstrate that, even using a CNN for prediction with a basic architecture, we are able to accurately predict MPI data for individual states in Nigeria. The main hope is that this methodology can be extended to regions that do not have a formally associated MPI. Given more computational power and using satellite imagery for regions that go beyond national borders as training data, there is potential to predict MPI for more granular regions (below even the state level) or in regions of the world the OPHI has yet to assign accurate MPIs.

Conclusions

To the best of our knowledge, this work makes the following key contributions:

- We curate a new dataset of publicly available, cloudless satellite imagery for all of Nigeria.
- We lay the first stone for utilizing deep learning methods for MPI prediction and metrics that go beyond single indices such as consumption expenditure or wealth.
- We build a custom CNN model that accurately predicts poverty directly from lower resolution (10m) satellite imagery, requiring only 0.01% of total land area, significantly reducing the amount of data and computational power required.
- Our results confirm that it is possible to predict MPI using samples of satellite imagery from even basic deep learning image models. Moreover, the more granular data we have regarding the components of MPI, the better our results. This suggests that satellite imagery could be a viable and cheaper method for estimating MPI data compared to current on-the-ground data collection methods.

Further work includes investigation of different sampling methods of satellite imagery, understanding the variation in prediction performance levels for the different components of MPI, and, if more computational resources are available, extending results to larger number of samples and larger datasets of MPI. In addition, there are other poverty metrics that are useful for prediction, and it is worth finding out which ones can be estimated cheaply via satellite imagery.

Acknowledgements

We thank Professor Milind Tambe and teaching fellows Paula Rodriguez Diaz and Sonja Johnson-Yu of CS 288 for their insights and guidance.

References

Abelson, B.; Varshney, K. R.; and Sun, J. 2014. Targeting direct cash transfers to the extremely poor. In *The 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-14)*.

Atia, M. 2014. Poverty Mapping.

Ayush, K.; Uzkent, B.; Tanmay, K.; Burke, M.; Lobell, D.; and Ermon, S. 2021. Efficient Poverty Mapping from High Resolution Remote Sensing Images. In *The Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21)*.

Babenko, B.; Hersh, J.; Newhouse, D.; Ramakrishnan, A.; and Swartz, T. 2017. Poverty Mapping Using Convolutional Neural Networks Trained on High and Medium Resolution Satellite Images, With an Application in Mexico. *arXiv preprint arXiv:1711.06323*.

Blumenstock, J.; Cadamuro, G.; and On, R. 2015. Predicting poverty and wealth from mobile phone metadata. *Science*, 350(6264): 1073–1076.

Engstrom, R.; Hersh, J. S.; and Newhouse, D. L. 2017. Poverty from Space: Using High-Resolution Satellite Imagery for Estimating Economic Well-Being. *World Bank Policy Research Working Paper No. 8284*.

Google Earth Engine. 2022. A planetary-scale platform for Earth science data analysis.

Humanitarian Data Exchange. 2021. Nigeria - Global Multidimensional Poverty Index (MPI). https://data.humdata.org/dataset/nigeria-mpi. Accessed: 2022-12-03.

Jean, N.; Burke, M.; Xie, M.; Davis, W. M.; Lobell, D. B.; and Ermon, S. 2016. Combining satellite imagery and machine learning to predict poverty. *Science*, 353(6301): 790–794.

Kingma, D. P.; and Ba, J. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Pandey, S.; Agarwal, T.; and Krishnan, N. C. 2018. Multi-Task Deep Learning for Predicting Poverty From Satellite Images. In *The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18)*.

Perez, A.; Yeh, C.; Azzari, G.; Burke, M.; Lobell, D.; and Ermon, S. 2017. Poverty Prediction with Public Landsat 7 Satellite Imagery and Machine Learning. *arXiv preprint arXiv:1711.03654*.

Piaggesi, S.; Gauvin, L.; Tizzoni, M.; Adler, N.; Verhulst, S.; Young, A.; Price, R.; Ferres, L.; Cattuto, C.; and Panisson, A. 2019. Predicting City Poverty Using Satellite Imagery. In *2019 Conference on Computer Vision and Pattern Recognition (CVPR-19) Workshop*.

Poverty, O.; and Initiative, H. D. 2022a. Global Multidimensional Poverty Index.

Poverty, O.; and Initiative, H. D. 2022b. Why map poverty?

United Nations Development Programme. 2020. 2020 Global Multidimensional Poverty Index (MPI): Charting pathways out of multidimensional poverty: Achieving the SDGs.

World Bank. 2022. Extreme poverty, 2015-2022.

Yeh, C.; Perez, A.; Driscoll, A.; Azzari, G.; Tang, Z.; Lobell, D.; Ermon, S.; and Burke, M. 2020. Using publicly available satellite imagery and deep learning to understand economic well-being in Africa. *Nature Communications*, 11(1): 1–11.

Appendices

Model Details

We utilize a basic convolutional neural network (CNN) with the following simple architecture: a convolutional layer with 1 input channel and 8 output channels, kernel size of 4, and stride of 1, followed by a max-pool layer of kernel size 4 and stride 2, followed by a second convolutional layer with 8 input channels and 16 output channels and kernel size of 2 and stride of 2, followed by a second max pool layer of kernel size 4 and stride of 2, followed by a dense layer going from 5184 inputs to 10 outputs, and a final dense layer condensing 10 inputs to one output. We train the CNN with mean squared error (MSE) loss with the Adam optimizer (Kingma and Ba 2014) with learning rate 2×10^{-5} for 100 epochs.