Mapping Poverty in the Philippines Using Machine Learning, Satellite Imagery, and Crowd-sourced Geospatial Information

Isabelle Tingzon [*1] **Ardie Orden** [*1] **Stephanie Sy** [1] **Vedran Sekara** [2] **Ingmar Weber** [3]
Masoomali Fatehkia [3] **Manuel Garcia Herranz** [2] **Dohyung Kim** [2]

Abstract

Mapping the distribution of poverty in developing countries is essential for humanitarian organizations and policymakers to formulate targeted programs and aid. However, traditional methods for obtaining socioeconomic data can be time-consuming, expensive, labor-intensive. Recent studies have demonstrated the effectiveness of combining machine learning with satellite images to estimate wealth in sub-Saharan African countries (Xie et al., 2015; Jean et al., 2016). In this study, we investigate the extent to which this method can be applied in the context of the Philippine archipelago to predict four different socioeconomic indicators: wealth level, years of education, access to electricity, and access to water. We also propose a cost-effective approach that leverages a combination of volunteered geographic information from OpenStreetMap (OSM) and nighttime lights satellite imagery for estimating socioeconomic indicators. The best models, which incorporate regional indicators as predictors, explain approximately 63% of the variation in asset-based wealth. Our findings also indicate that models trained on publicly available, volunteer-curated geographic data achieve the same predictive performance as that of models trained using proprietary satellite images.

1. Introduction

Despite best efforts in implementing poverty alleviation programs, the Philippines still lags behind its Southeast Asian neighbors in terms of poverty eradication, with approximately 22 million Filipinos living below the national poverty line (Philippine Statistics Authority PSA, 2018). A major challenge in fighting poverty today is the lack of reliable socioeconomic data, which is often expensive, time-consuming, and labor-intensive to collect. Conducting on-the-ground household surveys in the Philippines can cost up to 1.5M USD. Such surveys are done only every 3 to 5 years and are often aggregated to the regional or provincial level when reported to the public (Juan-Albacea, 2009). Without more granular and up-to-date data to guide their policies and programs, development organizations and government agencies risk allocating their limited resources in the wrong areas.

In recent years, major advancements in computer vision research and an increasing availability of geospatial resources have enabled novel methods for estimating socioeconomic indicators (Jean et al., 2016; Babenko et al., 2017; Engstrom et al., 2017). To tackle the problem of poverty eradication, we look towards combining machine learning with geospatial information as a fast, low-cost, and scalable means of providing granular poverty estimates. In this study, we examine the extent to which geospatial data including nighttime lights, daytime satellite imagery, human settlement data, and crowd-sourced information can be used to estimate socioeconomic well-being in the Philippines.

To summarize, our work primarily seeks to answer the following questions: (1) Are satellite-based methods developed for poverty prediction in other countries applicable within the Philippine context? and (2) How well do predictive models trained on publicly available crowd-sourced geospatial information compare against state-of-the-art satellite-based methods for Philippine poverty estimation?

2. Data and Pre-processing

2.1. Demographic and Health Survey

We used the 2017 Philippine Demographic and Health Survey (DHS) as a measure of ground truth for the socioeconomic indicators. Conducted every 3 to 5 years, the Philippine Statistical Authority (PSA) collects nationally representative information on social, economic, and health-

[*]Equal contribution [1]Thinking Machines Data Science, Manila, Philippines [2]UNICEF, New York, USA [3]Qatar Computing Research Institute, Doha, Qatar. Correspondence to: Isabelle Tingzon <issa@thinkingmachin.es>, Ardie Orden <ardie@thinkingmachin.es>.

Appearing at the International Conference on Machine Learning AI for Social Good Workshop, Long Beach, United States, 2019.

related outcomes across hundreds of households, which are grouped into clusters of 2 to 44 households (Philippine Statistics Authority PSA, 2018; Burgert et al., 2013). In line with the Sustainable Development Goals, we focused our analysis on a subset of survey questions and derived the following socioeconomic indicators from the DHS dataset:

Wealth Index. Our primary measure of socioeconomic well-being is the "wealth index", which is computed as the first principal component of attributes related to common asset ownership (e.g., roof material, television, housing material) on a per-household level. We get the mean wealth index per cluster as it is reported in the DHS dataset and do no further transformations.

Education completed. The DHS captures information on the number of years of education completed by household members over 6 years old. We aggregated this by computing the mean years of education completed across all households per cluster.

Access to Electricity. The DHS dataset contains information on the number of affirmative responses to the survey question related to access to electricity. We aggregated this information by getting the proportion of households with electricity access per cluster.

Access to Water. The DHS dataset contains information on the total travel time in minutes to get to a water source. If water source is on-site, time is set to zero. We get the mean time to access a water source across all households per cluster.

2.2. Nighttime Luminosity Data

The nighttime lights (NTL) data is taken from the Visible Infrared Imaging Radiometer Suite Day/Night Band (VIIRS DNB) for the year 2016, produced in 15 arc-second geographic grids (NOAA National Centers for Environmental Information, 2016). The VIIRS DNB NTL data includes a continuous luminosity level from 0 to 122 for the Philippines, with 0 being the darkest pixel. By observing the histogram of the nighttime light intensities, we assigned the nighttime light intensity values into the following five distinct classes: low intensity (zero pixel values), moderately low intensity (0.05-2), medium intensity (2-15), moderately high intensity (15-30), and high intensity (30-122).

2.3. Daytime Satellite Imagery

We retrieved a number of satellite images per cluster based on the cluster centroids reported in the DHS dataset, where each cluster location is defined by the mean latitude and longitude of the households, with added noise to preserve the privacy of the households (Philippine Statistics Authority PSA, 2018). Each cluster location is also labeled with a tag that indicates whether it is within a rural or urban area. We obtained up to 400 square tiles of satellite images within a 5 km radius for rural areas and up to 60 square tiles within a 2 km radius for urban areas. These tiles surround each cluster centroid and each tile corresponds to a pixel in the VIIRS DNB NTL dataset. Using Google Static Maps API, we downloaded a total of 134,540 images with a zoom level of 17, scale of 1, and pixel resolution of approximately 1.25 m. The size of each image is 400×400 pixels and matches the land area covered by a single pixel of night time lights data, which typically covers 0.25 km^2.

2.4. OpenStreetMap Data

More and more researchers are turning to volunteer-curated geographic information and open geospatial datasets to study socioeconomic development, social inequalities, and territorial conflicts (Gervasoni et al., 2018; Mahabir et al., 2018; Grippa et al., 2018). One of the more popular geospatial data crowd-sourcing platforms is OpenStreetMap (OSM), which contains global volunteered geospatial data. A recent study has found that user-generated road maps in OSM are approximately 83% complete as of 2016, with over 40% of countries having a fully mapped street network (Barrington-Leigh & Millard-Ball, 2017). We obtained OpenStreetMap (OSM) data for the Philippines from Geofabrik, an online repository for OSM data (Geofabrik GmbH, 2018). From this, we were able to extract extensive information related to the number of roads, buildings, and points of interests present within specified areas.

3. Methods

In this section, we describe the different methods used in predicting socioeconomic well-being. All models were evaluated using a five-fold nested cross validation scheme.

3.1. Satellite-based Transfer Learning Model

We implemented the satellite-based deep learning approach proposed by Xie et al. and later improved upon by Jean et al., with the assumption that nighttime lights act as a good proxy for economic activity (Xie et al., 2015; Jean et al., 2016; Mellander et al., 2015). As in Head et al., we began by fine-tuning a convolutional neural network (CNN) with VGG16 architecture that has been pre-trained on the ImageNet dataset to recognize 1000 different class labels (Krizhevsky et al., 2012; Head et al., 2017), with the goal of learning features that are useful for poverty prediction. We treat the problem as a classification task with five (5) night time intensity classes: low, moderately low, medium, moderately high, and high. We set aside 90% of the images for training and used the remaining 10% for the validation set. We dealt with the class imbalance by upsampling the minority classes (high, moderately high, and medium nighttime light intensities) and downsampling the low and moderately low light intensity classes to 30,000 images per class in order for all five classes to have the same number of training examples.

Like most models pretrained on ImageNet, the VGG16 model accepts 224 $\times$ 224 pixel images; meanwhile, our input images are 400 $\times$ 400 pixels. We proceeded to implement a fully convolutional architecture as described by Xie et al., which involves replacing the fully-connected top layers from the VGG16 model with randomly initialized fully convolutional top layers (Xie et al., 2015). This allows the model to accept input images of arbitrary sizes without losing information, unlike other techniques such as random cropping and scaling. Next, we augmented the training set using random horizontal mirroring, and used 50% dropout on the convolutional layers replacing the fully connected layers. We then began fine-tuning the full network using an Adam optimizer with an initial learning rate of 10^{-6} and a batch size of 32. We set the maximum number of epochs to 30, decreasing the learning rate by a factor of 10 whenever the validation loss began to plateau. We froze most of the layers, tuning only the last block of fully convolutional layers. At the end of training, we were able to achieve a 72% validation accuracy and 60% validation F1 score for the classification task.

For each image, we extract a 4,096-dimensional vector of activations in the top layer of the CNN, which are optimized to distinguish between different levels of night light luminosity. Each cluster has up to 400 images that we convert to feature vectors of learned representations; these feature vectors are then averaged into a single vector. Finally, we used these cluster-level feature vectors as input to a secondary regression model to predict the socioeconomic indicators. As in previous studies (Jean et al., 2016; Head et al., 2017), we used a ridge regression model to learn the mapping from cluster-level feature vectors to socioeconomic indicators.

3.2. OpenStreetMap Model

For each cluster, we extracted three types of OSM features, namely roads, buildings, and points of interest (POIs). These OSM features were extracted within a 5 km radius for rural areas and 2 km radius for urban areas, with each area centered on the cluster locations. We identified five types of roads in the dataset: primary, trunk, paved, unpaved, and intersection. In engineering road features, we followed the pre-processing technique described by Zhao and Kusumaputri, i.e., for each type of road, we calculated the distance to the closest road, total number of roads, and total road length per cluster (Zhao & Kusumaputri, 2016).

We also identified six different types of buildings: residential, damaged, commercial, industrial, education, health. For each type, we calculated the total number of buildings, the total area of buildings, the mean area of buildings, and the proportion of the cluster area occupied by the buildings. Finally, we identified over 100 different points of interests; for each cluster, we obtained the total number of each POI within a proximity of the area, e.g., number of banks, bars, cinemas, colleges, hotels, parks, etc.

We compared the performances of random forest regression models trained on the different types of OSM features, both separately and combined, for predicting socioeconomic well-being. Furthermore, we also conducted a series of experiments to determine the predictive performance of models trained using multiple data sources, with the hypothesis that using features from mixed data sources will bolster model performance. Specifically, we trained random forest regression models using a combination of OSM data and nighttime lights-derived features as input. Nighttime light features consist of summary statistics and histogram-based features, including the mean, median, maximum, minimum, covariance, skewness, and kurtosis, of the nightlight luminosity pixels within each cluster.

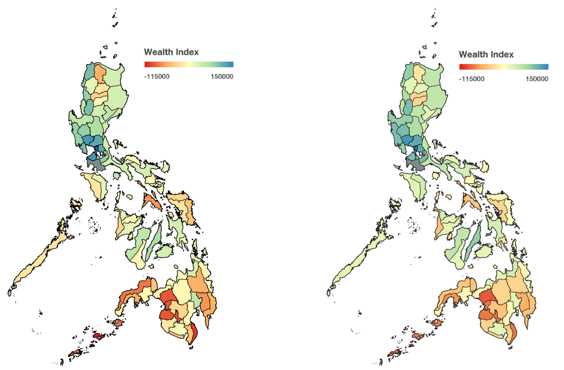
To our knowledge, this is the first paper to study multi-dimensional poverty prediction using a combination of crowd-sourced geospatial data and satellite data in the unique context of a developing nation in Southeast Asia.

4. Results and Discussion

4.1. Poverty Prediction using Satellite Imagery and Transfer Learning

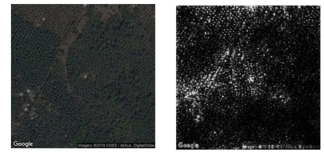
Past studies have published results on using deep learning methods for predicting wealth in sub-Saharan African countries (Jean et al., 2016) as well as non-African countries (Head et al., 2017). Predictive models achieved r-squared results ranging from 0.51 to 0.75 (Haiti: 0.51; Malawi: 0.55; Tanzania: 0.57; Nepal: 0.64 Nigeria: 0.68; Uganda: 0.69; Rwanda: 0.75). In this study, we tested how well the satellite-based deep learning approach performs in the Philippine setting.

Note that the Philippines, being an archipelago that consists of over 7,000 islands, required additional pre-processing steps in order to reduce noise in the dataset. Specifically, we removed satellite images with centroids located in bodies of water as well as images containing no human settlements using the High Resolution Settlement Layer (HRSL) developed by Facebook Research (Tiecke et al., 2017); by doing so we were able to see a notable rise in the r-squared score from 0.56 to 0.59 for wealth prediction. By increasing the number of nighttime light bins from the standard 3 to 5 and incorporating binary regional indicators as predictors, we were able to further improve the wealth index r-squared score to 0.63. As proof of concept, we show in Figure 1 a reconstruction of provincial-level poverty maps by aggregating cluster-level wealth estimates.



(a) Ground-truth Wealth Index (b) Predicted Wealth Index

Figure 1. Ground-truth wealth indices and cross-validated Philippine poverty predictions using the satellite-based transfer learning model aggregated to the provincial level.



(a) Low Nighttime Light Intensity Class

(b) Medium Nighttime Light Intensity Class

(c) Moderately High Nighttime Light Intensity Class

Figure 2. Class saliency maps for test satellite images in the nighttime light classification task. Maps were generated using a single back-propagation pass through the trained CNN.

Our findings also indicate that the method does not generalize for other socioeconomic indicators with the same level of accuracy as wealth in terms of r-squared (education: 0.47, access to electricity: 0.27, access to water: 0.10). We note that these results are consistent with the conclusions reached in Head et al., which states that high model performance cannot be expected when there is no clear relationship between the development indicator and nighttime lights (Head et al., 2017).

4.1.1. Visualizing Nighttime Light Classification Model

To visualize the nighttime light classification model, we generate class saliency maps based on a given image and class (Simonyan et al., 2013). We see in Figure 2 that the model identifies pixels related to roads and buildings as important for classifying medium to high nighttime light intensity; whereas, pixels related to trees and crops are given more weight when predicting low nighttime light intensity classes.

4.2. Poverty Prediction using Crowd-sourced Geospatial Information

We trained separate random forest regression models for each type of OSM feature (road, building, and POI). We found that using roads, buildings, or points of interests alone already explain 49-55% of the variance, with roads being the best predictor ($R^2 : 0.55$). Training a model using a combination of all three types of OSM features results in a higher r-squared (0.59). Furthermore, by combining OSM features with nighttime lights data and binary regional indicators, we were able to obtain an r-squared of 0.63 for wealth prediction. The r-squared results for education, electricity access, water access are 0.49, 0.36, and 0.09, respectively. Since our poverty prediction approach was optimized for predicting asset-based wealth, a more indicator-specific feature engineering and feature selection process may likely bolster performance.

We find that the performance of the the OSM-nightlights hybrid model (R^2: 0.63) compares similarly with the results of state-of-the-art satellite-based transfer learning model (R^2: 0.63). However, unlike satellite images from Google Static Maps which are proprietary and limited by licensing terms, both OSM and NTL data are publicly available and freely redistributable, making them an inexpensive alternative to daytime satellite images, which costs roughly 3,000 USD to acquire in order to generate granular poverty maps for the entire Philippines.

5. Conclusions

In this study, we implemented the satellite-based deep learning approach described by Xie et al. and Jean et al. in the Philippine setting (Xie et al., 2015; Jean et al., 2016). Our results confirm the applicability of the methodology, with the best model achieving an r-squared of 0.63 for estimating asset-based wealth. Moreover, this study demonstrates that the method cannot be trivially applied without taking into account the unique geography of the Philippine archipelago.

We also proposed an alternative cost-effective approach to poverty prediction that uses free and publicly available crowd-sourced geospatial information. Our findings indicate that a model trained on a combination of OSM and NTL-derived features also achieves an R^2 of 0.63. We conclude that both satellite images and volunteered geographic information are valuable tools for high resolution, real-time poverty mapping. Efforts in poverty mapping have great potential to help governments and humanitarian organizations better understand the spatial distribution of poverty and implement more evidence-based targeted interventions in developing countries.

6. Acknowledgments

This work was supported by the UNICEF Innovation Fund. We would also like to thank Neal Jean, Pia Faustino, Ram De Guzman, and Priscilla Moraes for the insightful discussions and valuable guidance.

References

Babenko, B., Hersh, J., Newhouse, D., Ramakrishnan, A., and Swartz, T. Poverty mapping using convolutional neural networks trained on high and medium resolution satellite images, with an application in mexico. *arXiv preprint arXiv:1711.06323*, 2017.

Barrington-Leigh, C. and Millard-Ball, A. The world's user-generated road map is more than 80% complete. *PLoS one*, 12(8):e0180698, 2017.

Burgert, C. R., Colston, J., Roy, T., and Zachary, B. *DHS Spatial Analysis Reports No. 7*. ICF International, Maryland, USA, 2013.

Engstrom, R., Hersh, J., and Newhouse, D. Poverty from space: using high-resolution satellite imagery for estimating economic well-being, 2017.

Geofabrik GmbH. Philippines - Geofabrik Download Server. https://download.geofabrik.de/asia/philippines.html, 2018.

Gervasoni, L., Fenet, S., Perrier, R., and Sturm, P. Convolutional neural networks for disaggregated population mapping using open data. In *DSAA 2018-5th IEEE International Conference on Data Science and Advanced Analytics*, pp. 1–10, 2018.

Grippa, T., Georganos, S., Zarougui, S., Bognounou, P., Diboulo, E., Forget, Y., Lennert, M., Vanhuysse, S., Mboga, N., and Wolff, E. Mapping urban land use at street block level using openstreetmap, remote sensing data, and spatial metrics. *ISPRS International Journal of Geo-Information*, 7(7):246, 2018.

Head, A., Manguin, M., Tran, N., and Blumenstock, J. E. Can human development be measured with satellite imagery? In *Proceedings of the Ninth International Conference on Information and Communication Technologies and Development*, pp. 8. ACM, 2017.

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

Juan-Albacea, Z. V. Small area estimation of poverty statistics. Technical report, Citeseer, 2009.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pp. 1097–1105, 2012.

Mahabir, R., Croitoru, A., Crooks, A., Agouris, P., and Stefanidis, A. News coverage, digital activism, and geographical saliency: A case study of refugee camps and volunteered geographical information. *PloS one*, 13(11): e0206825, 2018.

Mellander, C., Lobo, J., Stolarick, K., and Matheson, Z. Night-time light data: A good proxy measure for economic activity? *PloS one*, 10(10):e0139779, 2015.

NOAA National Centers for Environmental Information. VIIRS DNB Dataset. https://ngdc.noaa.gov/eog/viirs/download_dnb_composites.html, 2016. Online; accessed 18 December 2018.

Philippine Statistics Authority PSA, I. Philippines national demographic and health survey 2017, 2018. URL http://dhsprogram.com/pubs/pdf/FR347/FR347.pdf.

Simonyan, K., Vedaldi, A., and Zisserman, A. Deep inside convolutional networks: Visualising image classification models and saliency maps. *arXiv preprint arXiv:1312.6034*, 2013.

Tiecke, T. G., Liu, X., Zhang, A., Gros, A., Li, N., Yetman, G., Kilic, T., Murray, S., Blankespoor, B., Prydz, E. B., and Dang, H. H. Mapping the world population one building at a time. *CoRR*, abs/1712.05839, 2017. URL http://arxiv.org/abs/1712.05839.

Xie, M., Jean, N., Burke, M., Lobell, D., and Ermon, S. Transfer learning from deep features for remote sensing and poverty mapping. *arXiv preprint arXiv:1510.00098*, 2015.

Zhao, L. and Kusumaputri, P. Openstreetmap road network analysis for poverty mapping, 2016.