Measuring Poverty in Vietnam with Machine Learning and Satellite Images

Dang Ngoc Huy
h.dang@mpp.hertie-school.org

Abstract

Mapping poverty to better target aid and development is a difficult business for a number of reasons: the reliance on the collection of household data which is time-consuming and costly; household surveys and data are often not available for many countries of interest and when they are, it is usually not at the desired frequency and this time-lag creates problems in the decision making process for development and aid agencies; to get a sense of a region's current development, it is often essential to use predictions to fill in the gap in the time-series data of that region's economic indicators, which can be questionable and unreliable. Recent researches and development in the field of computer vision and deep learning have displayed the effectiveness of employing publicly available satellite images to map out impoverished areas in the sub-Saharan African region (Neal Jean et al., 2016). This studies seeks to investigate the extent to which this methodology can be applied in the context of Vietnam to predict different socioeconomic and poverty indicators...

1. Introduction

In 2015, the United Nations adopted 17 Sustainable Development Goals, one of which is the complete eradication of poverty by 2030. To realize such an ambition, the first step is essentially to identify the areas most affected by poverty, to better target aid and development programs. However, mapping poverty is a costly endeavor with complex politics, limited and unreliable availability of information. To provide a detailed picture of economic-wellbeing of a region, oftentimes an elaborate and comprehensive household survey is needed which prompts a series of concerns from funding, cost effectiveness to governments wanting to shield their inadequate engagement with their poorest citizens.

The use of publicly available satellite imagery, combined with computer vision and deep learning, on the other hand, is relatively straightforward, simple and low cost, while bypassing the issue of institutional complication, save for a few issues regarding personal privacy of affected individuals. Furthermore, the challenges and rewards of understanding poverty on a granular level from images in space are indeed a scientific endeavor worthy of pursuing. A number of researches have been conducted in this promising field to gauge the effectiveness of this new paradigm in extracting information from satellite images in helping poverty eradication, with successes as well as limitations. This studies has been implemented with the hope of building up on what has been done thus far and providing another narrative of whether this methodology could achieve a favorable outcome in a completely different country and setting such as Vietnam.

This paper will first provide an overview of the current literature in the area of poverty measurement with the use of satellite images analysis. Then, the proposed method relevant to the research will be outlined along with several concerns regarding its implementation as well as a brief overview of poverty and inequality in Vietnam to provide the basis of ground truth for analysis. Finally, an elaboration of the experiments themselves followed by an analysis of the results and concluding remarks will be detailed.

2. Related Work

To resolve the problem of poverty identification in the scarcity of available data, Neal Jean and his team of researchers at Stanford's University's Department of Earth System Science proposed the novel approach of analyzing satellite images of the planet's surface to pinpoint regions and cluster areas that might be regarded as ravaged by poverty in their seminal work: *Combining satellite imagery and machine learning to predict poverty*, Science, vol. 353, no. 6301, pp. 790–794, 2016.

To train a model that can identify impoverished areas, Jean and his teams utilized three different data types: images of luminosity at night time which could signal economic level; images of daytime from which useful information on landscape features and socioeconomic data could be extracted; and finally survey data on household consumption expenditure and asset wealth as validation for the findings extrapolated from the analysis of satellite images.

A number of challenges plague this pursuit. The theoretical framework employed by the authors rests on the assumption that there is a correlation between the nightlight luminosity of a region reflected in their satellite images and their levels of economic activities since less intensity could potentially indicate less consumption and wealth level in an area. However, as the authors observed, this method is unable to detect granular economic activities in regions that are near and below the international poverty line, which makes it less useful to distinguish the poverty-striken areas of interest. In addition, satellite images are notoriously unstructured, with little to no labels available, which makes supervised learning to classify and extract features difficult to implement.

To accommodate for these drawbacks, the learning process of their research occurred in three consecutive phases, each of which supplemented one another to improve the tracking of clusters of people living in poverty: in the first phase, a transfer learning approach is employed with a convolutional neural network (CNN) model that has been trained on ImageNet to detect features such as edges and corners; in the second phase, this CNN model is further trained to estimate nighttime luminosity level based on the daytime satellite images on a global scale, with the idea that the model can discern certain features in the daylight images such as segmentations of land use, signs of human assets and activities as well as other useful geographic information that are indicative of the variation in nightlight luminosity - an acceptable proxy for economic activities; in the final phase, economic survey data on clusters of households is combined with the features extracted by the neural network to train and learn the patterns and connection between the different features that it had identified from the images and the household spending and asset possession using a ridge regression model, ultimately to estimate and forecast the poverty indicators of interest. Focusing on five countries from Africa - Nigeria, Tanzania, Uganda, Malawi, and Rwanda, the researchers were able to explain up to 75 percent of the variation in economic outcomes at the local level in cross-validated sets just based on the satellite images which were freely available in the public domain.

One particular ingenious implementation is the analysis of nightlight luminosity and how it adds to the ability of the model to extract information. Although nightlight information is an flawed proxy for economic levels since it demonstrated little correlation with the variation for expenditure spending at a lower level, daytime information and the socioeconomic information that can be extracted such as materials used on roof of housing to distance from the cluster of households of interest to urban areas are good estimator for the expenditure even at a granular level. Therefore, by training the CNN architecture on both information in nightlight and daytime imagery, the model can learn and identify features that can capture variations across the entire expenditure and household consumption distribution, leading to a more accurate picture of areas affected by poverty and of lower economic-well-being in comparison to others. In this approach of training on daytime satellite images to estimate nightlight luminosity, the model can learn to identify landscape features without needing explicit annotations and labelling, reducing the need for direct supervision while still able to detect useful information relevant to the research question.

With a model that now has an understanding of the correlation between the features present in the satellite images and poverty level, the research could then map out predictions on regions where there are gaps in the survey data, helping to identify previously unknown areas for aid targeting.

Following a similar investigation, Babenko et al. also seeks to understand the usefulness of satellite image analysis with deep learning in identifying poverty through their paper *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. Using a different CNN model and architecture setup with GoogleNet, also finetuned from ImageNet, their research was able to successfully estimate land usage through analysing satellite images and subsequently utilize this as a predictor for poverty indicators. Their best model, thus, could explain up to 57 percent of the poverty variation in poverty in a sample set of 10 percent of municipalities in Mexico. Their conclusion, similar to Neal Jean et al., is that a CNN architecture learning from satellite images can be an end-to-end solution to understand and estimate poverty, in the face of unavailable survey data. However, they also underscored that the training process and how it influenced out of sample validation needs to be further researched and understood.

Another research that helps to solidify the confidence of this approach is a World Bank funded policy research paper from Ryan Engstrom, Jonathan Hersh and David Newhouse looking at how highly spatial satellite images can estimate poverty and economic level in Sri Lanka in *Poverty from Space Using High-Resolution Satellite Imagery for Estimating Economic Well-Being*, World Bank Research Group, Poverty and Equity Global Practice Group, December 2017. In a similar vein as previous studies, Engstrom et al. extracted features through the analysis of satellite view to identify landscape information such as agricultural practices, density of building and infrastructures, density and lengths of roads, the number of cars, roofing materials, as well as other textual details. A simple linear regression with these features as input as predictors was able to explain up to 60 percent of poverty headcount rates and average log consumption for 1,291 administrative units in Sri Lanka.

These researches and their successes have demonstrated that the use of satellite images and deep learning is certainly a promising avenue to help understand the distribution of wealth and poverty. However, their fluctuating results also call into question the reliability of the approach and how it can be translated to other settings. As any scientific endeavors, the more evidence there is, the more confident we can be of its effectiveness. This studies, therefore, seeks to supplement the credence that this approach can work in different setting and countries, focusing on Vietnam as a nation that has its fair share of poverty and inequality issues.