Evaluating the Relationship between Spatial and Spectral Features Derived from High Spatial Resolution Satellite Data and Urban Poverty in Colombo, Sri Lanka

Dr. Ryan Engstrom
Department of Geography
The George Washington University
Washington, District of Columbia
rengstro@gwu.edu

Dr. David Newhouse
Poverty Global Practice Group
The World Bank
Washington, District of Columbia
dnewhouse@worldbank.org

Vishwesh Haldavanekar
Data Science Program
The George Washington University
Washington, District of Columbia
vishwesh_s_h@gwmail.gwu.edu

Andrew Copenhaver
Department of Geography
The George Washington University
Washington, District of Columbia
agcopenhaver@gwmail.gwu.edu

Jonathan Hersh
Department of Economics
Boston University
Boston, Massachusetts
jhersh@bu.edu

***Abstract*— In this study we seek to map urban poverty in Colombo, Sri Lanka using spectral and spatial features estimated from high spatial resolution satellite imagery. For this study we calculated 165 spectral and spatial features at a block size of 16m and a range of scales, from three Quickbird scenes, collected in 2010 which cover 316 Grama Niladhari (GN) census units within the District of Colombo and includes the urban area of Colombo, Sri Lanka. The features calculated include linear support regions (LSR), linear binary pattern moments (LBPM), PanTex, Histogram of Oriented Gradients (HoG), Speeded Up Robust Features (SURF), Fourier Transform (FT), Gabor, the mean of each of the blue, green, red, and near-infrared spectral bands, as well as the Normalized Difference Vegetation Index (NDVI). For each GN census unit (avg. size of 2.17 sq. km), the zonal sum, mean, and standard deviation of all 165 features were calculated. For each GN unit, the 10/20/30/40th percentiles of the national distribution of household estimates of predicted per capita consumption were calculated using data from the 2011 Sri Lankan Census to provide an estimate of poverty. Results indicate that the combined spatial and spectral features were able to explain up to 54% of the variation in poverty when using a simple, ordinary least squares linear regression model.**

Keywords—poverty estimates; texture; spatial feature extraction; urban; census; multiple regression

I. Introduction

Mapping poverty has traditionally been performed by either using household survey data or by combining these results with census data collected by national statistics offices. This is a time consuming, labor intensive, and costly process that limits the ability to collect data in many areas, with many countries never even collecting the data [1]. Moreover, the number of samples needed when conducting only a household survey leads them to only be valid across large spatial areas and when combining them with census data to produce small area estimates leads them to be dependent on these data which are only captured every 10 years at the most. One potential way to improve upon poverty estimation is to use data sources, such as those derived from satellite imagery, which has been shown to be useful in mapping local variations in poverty [2].

With the improved spatial resolution of satellite imagery, the mapping of poverty within urban areas has been steadily increasing in the last decade with a general focus on the mapping of the presence/absence of slums [3]. As part of this research theme a remote sensing scene model has been developed that has helped to discern what a slum looks like in imagery [4]. While this scene model helps to understand what a slum might look like within imagery, simply visually interpreting slum versus non-slum areas in high spatial resolution satellite imagery can be difficult due to varying local definitions of slums [4] as well as differing morphologies of slums in different locales [5]. As a result, developing methods for automated mapping of these areas using remote sensing has proven to be a challenging task [3, 4, 6, 7, 8].

Another way to look at variability in poverty within a city is to map it directly, as estimated through census and/or survey data.

The recent work published by Jean et al. [9] indicates that satellite data derived values can be used to map variations in poverty. Their research utilized a convolutional neural network (CNN); however, as CNNs require an enormous amount of training data [10], their model was trained using a multistep approach that they term "transfer learning". Generally, their transfer learning approach tuned an existing CNN to extract features, which summarize the relationship between nighttime lights intensities (a general predictor of economic activity) and high resolution, daytime satellite imagery. These features were then used to train ridge regression models that were used to predict household expenditures and assets at a scale roughly equivalent to that of a village. Their results are encouraging, being able to explain up to 55% of variation in average consumption and 75% of variation in average asset wealth. However, as was stated above, the difficulty in training a CNN makes this approach particularly challenging and their out of sample estimates are much lower than their in sample results.

Previous research has indicated that several simpler, traditional spectral and spatial features extracted from high resolution imagery are related to numerous potential indicators of poverty, such as informal housing, housing quality, population density, lack of solid waste collection, and lack of improved sanitation [11, 12, 13]. Extracting these spatial features is an image processing technique, which derives information related to spatial patterning, structure, orientation, texture, and irregularity from windows of image pixels [14]. That spatial and spectral features are related to indicators of poverty is relatively intuitive, as urban morphology and structure are related to socioeconomic conditions [15], and spatial and spectral features extract information related to the structure and morphology of an area. By deriving a number of spatial and spectral features from high resolution satellite imagery, including: Fourier Transform (FT) [16], Histogram of Oriented Gradients (HoG) [17], Local Binary Patterns (LBP) [18], Line Support Regions (LSR) [19], a texture derived built-up presence index (PanTex) [20], Speeded Up Robust Features (SURF) [21], Gabor Filters [22], the individual and combined spectral band means and the Normalized Difference Vegetation Index (NDVI) [23], this study investigates the ability to directly map poverty estimated at different levels using a combination of several features within a multiple, ordinary least squares (OLS) regression model for the urban area of Colombo, Sri Lanka.

II. METHODS

A. Data

Three, cloud free, multispectral (Blue, Green, Red, and NIR) Quickbird scene sections, captured in 2010 were obtained from LAND INFO Worldwide Mapping, LLC (LAND INFO). These images were orthorectified by LAND INFO to ensure they matched the census data and re-sampled to 2m spatial resolution. These images cover an approximately 348 square kilometer area, spanning the city of Colombo, Sri Lanka and the surrounding areas. Poverty statistics at the GN level were estimated using household level estimates of poverty derived from the 2011 Sri Lankan Census of Population and Housing. The method and models are identical to those used to generate official poverty estimates at the DS Division level from the census [24]. For each household in the census, per capita consumption was estimated based on models developed from the 2012-2013 Household Income and Expenditure Survey (HIES). The models, which were estimated using the PovMap 2.0 software package, employ the Elbers, Lanjouw, and Lanjouw [25] methodology. Per capita consumption in the HIES was deflated using a district-level spatial price index derived from unit value food prices for households in between the 10th and 40th percentiles of per capita consumption [26]. We derive GN headcount poverty statistics using the Foster-Greer-Thorbecke methodology [27] at four different thresholds of poverty -- 10/20/30/40% of national income -- according to the formula:

$$povrate_{j,z_k} = \frac{1}{N_j * 100} \sum_{i=1}^{N_j} \sum_{s=1}^{100} I(y_{is} < Z_k) \left(\frac{Z_k - y_{is}}{z_k} \right)^0 = \frac{1}{N_j * 100} \sum_{i=1}^{N_j} \sum_{S=1}^{100} I(y_{is} < Z_K)$$

where $povrate_{j,z_k}$ is the poverty headcount rate in GN Division j at the z_k poverty threshold, defined at 10/20/30/40 percent of national income. The variable y_i is per capita household consumption for individual i, averaged over the 100 simulations, N_j is the total population of GN j, S indexes simulations, each of which contains a random draw of the error term. We calculated the national poverty threshold level, z_k, as the $k \in \{10,20,30,40\}$ percentile of the national predicted per capita consumption in the census averaged across the 100 simulations. I is an indicator function that in this case takes on one if individual i's per capita household consumption lies below the relative poverty threshold Z_k. Since only the headcount rate was considered, this is equivalent to the ratio of the number of individuals below the threshold to the total population in each GN Division. The total sample for this study contains 296 GN divisions, with each GN division containing an average of 10,000 persons according to the 2011 Census, and covering an average area of 2.12 square kilometers. The GNs used in this study cover approximately 2.14 square kilometers (Fig. 1).

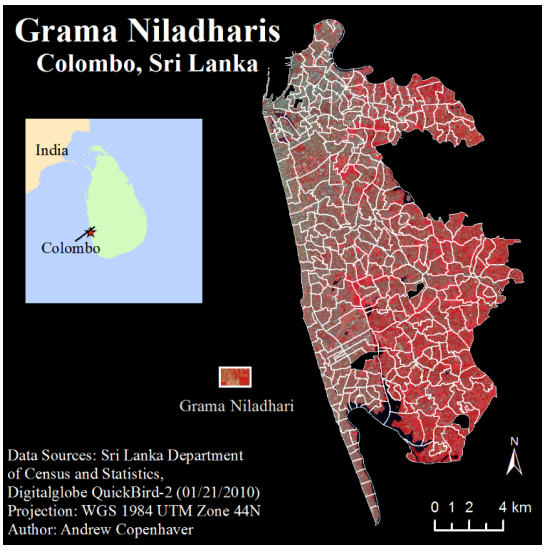

Fig. 1. Study area map displaying imagery and Grama Niladhari (GN) divisions used.

B. Spatial and Spectral Feature Calculation

Spatial and spectral features were created using a block size of eight pixels and scales of 8, 16, 32, and 64 meters using a methodology similar to Graesser et al. [11]. Block size can be thought of as the number of input pixels that are aggregated to create one output pixel, while the scale size is the size in input image pixels of the moving window used to calculate the features. The seven textural features calculated for this study were: Histogram of oriented gradients (HoG), which captures edge orientations and sorts them into a histogram [17], PanTex, which is a built-up presence index derived from the grey-level co-occurrence matrix (GLCM) [20], Line support regions (LSR), which characterize line attributes [28], Local binary pattern moments (LBPM), which define contiguous regions of pixel groups and sorts them into a histogram [29], Fourier transform (FT) which examines pattern frequency across an image [30], Gabor, a linear Gaussian filter used for edge detection [22], Speeded Up Robust Features (SURF), an algorithm that extracts key points (i.e., edges and corners) from an image through pyramidal Gaussian based decomposition [21], the Normalized Difference Vegetation Index (NDVI), the most widely used vegetation index that provides information about the presence and abundance of vegetation [23], the mean of the combined spectral bands, and the means of the four individual bands, Blue, Green, and Near Infrared. The feature extraction process resulted in an output image comprised of 165 bands at a spatial resolution of 16m, as the Quickbird scenes were re-sampled to two meters during orthorectification.

Finally, in order to place the extracted spatial and spectral features on the same scale as the GN level estimates of poverty, a polygon shapefile, which contained the GN Division boundaries was used to calculate zonal statistics (mean, standard deviation, and sum) for each of the 165 bands of the output image.

III. Regression analysis

Due to the high dimensionality of the zonal statistics output (456 output variables) and the fact that many of the output variables exhibited high levels of multicollinearity, traditional approaches to variable selection, such as stepwise methods proved inadequate. As a result, we developed a hybrid methodology for variable selection, which first created three different regressions (Partial Least Squares, Ridge and LASSO) using all of the spatial and spectral features as independent variables and predicted poverty as the dependent variable. Next, we took the top 20 variables (based on the regression coefficients) from each of the three regressions and performed a Shapley value analysis [31]. Using a Shapley value threshold of 1% we eliminated low contribution variables from the model. This yielded 38 variables, which were then used within an OLS regression. Variables were then interactively removed from the OLS regression based on their significance, their relative importance to the adjusted R^2 value as well as their variance inflation factors (VIF). A VIF of 7.5 was the threshold at which variables were excluded from the models.

In order to check for spatial autocorrelation in the standardized residuals, which indicates that potentially important independent variables are missing from the model, a global Moran's I test was run at a fixed neighborhood distance of 4.965 meters, which was the distance determined to include at least eight neighboring features for all observations. A global Moran's I z-score greater than 2.58 indicates that residuals are clustered.

Finally, to determine the percent R^2 contribution for each variable included within the final model, a Shapley decomposition analyses [31] was performed.

IV. Results

Table 1 describes the variables that were included within the model based on the variable selection procedure.

TABLE 1. SUMMARY OF SELECTED FEATURES

Feature	*Block*	*Scale*	*Zonal Statistic*	*Shapley R^2 Contribution (%)*	*VIF*
Gabor Filter, 280º rotation	8	64	SD	4.79	4.59
Gabor Filter, 20 º rotation	8	8	SD	5.63	5.90
HoG, histogram variance	8	64	SD	8.63	1.49
HoG, histogram kurtosis	8	64	SD	4.53	2.46
HoG, histogram skew	8	8	SD	4.98	3.35
LBPM, histogram kurtosis	8	16	SD	6.84	2.96
LBPM, histogram mean	8	64	Mean	1.62	1.29
LBPM, histogram kurtosis	8	64	SD	2.55	1.60
LSR, mean line length	8	16	SD	3.80	3.85
LSR, line variance	8	64	SD	3.05	4.72
LSR, mean line length	8	8	Mean	2.52	2.47
LSR, line variance	8	8	SD	1.75	1.49
Combined Band Mean	8	64	Mean	21.00	5.55
NDVI	8	64	Mean	13.68	5.85
NDVI	8	8	Sum	2.38	3.30
SURF	8	32	SD	1.67	2.35
SURF	8	64	Sum	10.60	2.52

The linear regression analyses indicate that a relatively high percentage of the variation in poverty at multiple levels can be explained by spatial and spectral features. The coefficients of determination for each estimated poverty level are as follows:

- 10^{th} percentile: Adjusted $R^2 = 0.54$
- 20^{th} percentile: Adjusted $R^2 = 0.54$

- 30^{th} percentile: Adjusted $R^2 = 0.53$
- 40^{th} percentile: Adjusted $R^2 = 0.52$
- Average Simulated Household Consumption: Adjusted $R^2 = 0.46$

All regressions are significant at an alpha value of 0.01.

Finally, the results of the global Moran's I test indicated that statistically significant spatial autocorrelation is present in the standardized residuals for the model at each level of estimated poverty. Z-scores of 7.15, 9.76, 11.62, 13.37 and 17.40 were associated with the 10^{th} percentile, 20^{th} percentile, 30^{th} percentile, 40^{th} percentile and average simulated household consumption, respectively.

V. Discussion

These results indicate that spatial and spectral features perform reasonably well on their own at explaining poverty, with adjusted R^2 values ranging from 0.46 to 0.54. However, we also found spatial autocorrelation in the standardized residuals, which is indicative that potentially important explanatory variables are missing from the models. This is not especially surprising given the complexity of urban poverty and the fact that urban poverty is more than a simple precursor and/or result of the spatial arrangement of on-the-ground objects. Future research will investigate what other features may aide in explaining variations in poverty and how well these models can be transferred to other cities.

References

[1] Serajuddin, Umar, Hiroki Uematsu, Christina Wieser, Nobuo Yoshida, and Andrew Dabalen. "Data deprivation: another deprivation to end." *World Bank Policy Research Working Paper* 7252 (2015).

[2] Duque, Juan C., Jorge E. Patino, Luis A. Ruiz, and Josep E. Pardo-Pascual. "Measuring intra-urban poverty using land cover and texture metrics derived from remote sensing data." *Landscape and Urban Planning* 135 (2015): 11-21.

[3] Kuffer, Monika, Karin Pfeffer, and Richard Sliuzas. "Slums from Space—15 Years of Slum Mapping Using Remote Sensing." *Remote Sensing* 8, no. 6 (2016): 455.

[4] Kohli, Divyani, Alfred Stein, and Richard Sliuzas. "Uncertainty analysis for image interpretations of urban slums." *Computers, Environment and Urban Systems* 60 (2016): 37-49.

[5] Sliuzas, Richard, and Monika Kuffer. "Analysing the spatial heterogeneity of poverty using remote sensing: typology of poverty areas using selected RS based indicators." *Remote Sensing–New Challenges of High Resolution, Bochum* (2008): 5-7.

[6] Taubenböck, Hannes, and N. J. Kraff. "The physical face of slums: a structural comparison of slums in Mumbai, India, based on remotely sensed data." *Journal of Housing and the Built Environment* 29, no. 1 (2014): 15-38.

[7] Kohli, Divyani, Pankaj Warwadekar, Norman Kerle, Richard Sliuzas, and Alfred Stein. "Transferability of object-oriented image analysis methods for slum identification." *Remote Sensing* 5, no. 9 (2013): 4209-4228.

[8] Engstrom, Ryan, Avery Sandborn, Qin Yu, Jason Burgdorfer, Douglas Stow, John Weeks, and Jordan Graesser. "Mapping slums using spatial features in Accra, Ghana." In *2015 Joint Urban Remote Sensing Event (JURSE)*, pp. 1-4. IEEE, 2015.

[9] Jean, Neal, Marshall Burke, Michael Xie, W. Matthew Davis, David B. Lobell, and Stefano Ermon. "Combining satellite imagery and machine learning to predict poverty." *Science* 353, no. 6301 (2016): 790-794.

[10] Tajbakhsh, Nima, Jae Y. Shin, Suryakanth R. Gurudu, R. Todd Hurst, Christopher B. Kendall, Michael B. Gotway, and Jianming Liang. "Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?." *IEEE transactions on medical imaging* 35, no. 5 (2016): 1299-1312.

[11] Graesser, Jordan, Anil Cheriyadat, Ranga Raju Vatsavai, Varun Chandola, Jordan Long, and Eddie Bright. "Image based characterization of formal and informal neighborhoods in an urban landscape." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 5, no. 4 (2012): 1164-1176.

[12] Sandborn, Avery, and Ryan N. Engstrom. "Determining the relationship between census data and spatial features derived from high-resolution imagery in Accra, Ghana." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9, no. 5 (2016): 1970-1977.

[13] Engstrom, R., Copenhaver, A., and Qi, Y. "Evaluating the use of Multiple Imagery-Derived Spatial Features to Predict Census Demographic Variables in Accra, Ghana". In 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), In Press.

[14] Herold, Martin, XiaoHang Liu, and Keith C. Clarke. "Spatial metrics and image texture for mapping urban land use." *Photogrammetric Engineering & Remote Sensing* 69, no. 9 (2003): 991-1001.

[15] Taubenbock, H., M. Wurm, N. Setiadi, N. Gebert, A. Roth, G. Strunz, J. Birkmann, and S. Dech. "Integrating remote sensing and social science." In *2009 Joint Urban Remote Sensing Event*, pp. 1-7. IEEE, 2009.

[16] Steven, W. Smith. "The scientist and engineer's guide to digital signal processing." *California Technical Pub* (1997).

[17] Dalal, Navneet, and Bill Triggs. "Histograms of oriented gradients for human detection." In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1, pp. 886-893. IEEE, 2005.

[18] Ojala, Timo, Matti Pietikainen, and Topi Maenpaa. "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns." *IEEE Transactions on pattern analysis and machine intelligence* 24, no. 7 (2002): 971-987.

[19] Unsalan, C. "Gradient-magnitude-based support regions in structural land use classification." *IEEE Geoscience and Remote Sensing Letters* 3, no. 4 (2006): 546-550

[20] Pesaresi, Martino, Andrea Gerhardinger, and François Kayitakire. "A robust built-up area presence index by anisotropic rotation-invariant textural measure." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 1, no. 3 (2008): 180-192.

[21] Bay, Herbert, Tinne Tuytelaars, and Luc Van Gool. "Surf: Speeded up robust features." In *European conference on computer vision*, pp. 404-417. Springer Berlin Heidelberg, 2006.

[22] Gabor, Dennis. "Theory of communication. Part 1: The analysis of information." Electrical Engineers-Part III: Radio and Communication Engineering, Journal of the Institution of 93, no. 26 (1946): 429-441.

[23] Tucker, Compton J. "Red and photographic infrared linear combinations for monitoring vegetation." *Remote sensing of Environment* 8, no. 2 (1979): 127-150.

[24] Sri Lanka Department of Census and Statistics and the World Bank Group. "The Spatial Distribution of Poverty in Sri Lanka." (2015)

[25] Elbers, Chris, Jean O. Lanjouw, and Peter Lanjouw. "Micro–level estimation of poverty and inequality." *Econometrica* 71, no. 1 (2003): 355-364.

[26] Newhouse, David, Becerra Suarez, and Dung Doan. 2016. *Sri lanka poverty and welfare: Recent progress and remaining challenges*. The World Bank Group.

[27] .Foster, James, Joel Greer, and Erik Thorbecke. "A class of decomposable poverty measures." *Econometrica: Journal of the Econometric Society*(1984): 761-766.

[28] Yu, Won Pil, Gil Whoan Chu, and Myung Jin Chung. "A robust line extraction method by unsupervised line clustering." Pattern recognition 32, no. 4 (1999): 529-546.

[29] Wang, Li, and Dong-Chen He. "Texture classification using texture spectrum." Pattern Recognition 23, no. 8 (1990): 905-910.

[30] Smith, Steven W. "The scientist and engineer's guide to digital signal processing." (1997).

[31] Shorrocks, Anthony F. "Decomposition procedures for distributional analysis: a unified framework based on the Shapley value." Journal of Economic Inequality (2013): 1-28.