2019

IARIW-World Bank

Special IARIW-World Bank Conference "New Approaches to Defining and Measuring Poverty in a Growing World" Washington, DC, November 7-8, 2019

Using Satellite Data to Guide Urban Poverty Reduction

Thomas Pave Sohnesen

Peter Fisker

David Malmgren-Hansen

Paper Prepared for the IARIW-World Bank Conference

Washington, DC, November 7-8, 2019

Session 2B: Data Methods for Improved Poverty Measurement

Time: 14:15 – 16:15, September 7

Using Satellite Data to Guide Urban Poverty Reduction

Corresponding author:

Thomas Pave Sohnesen

World Bank

1818 H Street, NW Washington, DC 20433 USA

Co-authors:

Peter Fisker

Development Economics Research Group, University of Copenhagen

David Malmgren-Hansen

Technical University of Denmark

Abstract

Poverty reduction is increasingly an urban challenge, and a challenge that continues to be hampered by lack of data. One such example is the urban social safety net program implemented by the Government of Mozambique, that is spatial in nature, but works without any data on the within-city spatial distribution of poverty. The lack of detailed data on poverty is common in many developing as well as middle-income countries. This study applies Convolutional Neural Networks on high-resolution satellite images of cities in Mozambique, and combines their outputs with household level geo-referenced survey data. The results show that readily available data sources can generate detailed neighborhood-level poverty maps, providing key operational guidance for implementation of the urban social safety net. Importantly, the approach is highly automatic, applicable at scale, and cost-effective. It is thus a key step forward in the application of remote sensing image recognition for urban poverty reduction.

Keywords: Poverty, Social protection, Remote sensing, Convolutional neural networks, image recognition.

Introduction

As stated by the UN in their Sustainable Development Goals, poverty reduction is a key global challenge. Poverty reduction requires sustained efforts and often the combination of multiple policies. In the poorest countries, the challenge is especially difficult as policies are designed with very limited available data. The second-best solution is often to fill the information gap through household surveys. Unfortunately, such surveys are too expensive to be a general solution. The World Bank estimates that monitoring and tracking poverty in the poorest countries will cost USD 945 million between 2016 and 2030 (1), while they rarely provide current and continuous data. Further, such surveys normally estimate vital statistics at urban/rural and regional levels, while they provide no guidance on planning or differences within specific urban environments. This paper's approach provides key insights at within-city level, but also illustrates a potential for much wider use.

The rapid expansion of available high-resolution satellite images provides an opportunity to close some of this information and data gap. Satellite images of cities by themselves, as seen by the human eye, are often very informative of the general living situation in any given location. Human observations are an insufficient tool for policy development, but satellite images combined with image recognition and machine learning methods do have potential to reduce the data gap and thereby pave the road for more efficient anti-poverty policies. Recent progress in the use of images for prediction of poverty includes the use of nighttime light data and daytime images to estimate village poverty rates in five African countries(2). Object identification of buildings and their quality as well as cars, combined with geospatial data on roads and farmland, are used to predict municipal poverty rates in Sri Lanka (3). Similarly, object identification of water source, roof quality and lighting source are used to predict poverty at sub-district levels in Uttar Pradesh, India(4). Object identification of roof quality, but not poverty itself, also provides the basis of allocation of anti-poverty transfers in some villages in Uganda and Tanzania(5). Prediction of municipal poverty rates in Mexico, on the other hand, is done by relating poverty rates directly to images without specifying any objects (6). Even single households' poverty status are predicted using remote sensing data (7).

This study contributes to the nascent literature by estimating poverty scores for locations smaller than 115m x 115m, which is an improvement on most previous studies that estimate poverty rates of larger areas. To the authors' knowledge, this is also the first study that provides a comparison of the direct approach, relating poverty directly to images, and the indirect approach of prediction poverty based on identified objects correlated with poverty. Importantly, unlike the estimates at household level (7), that rely on manually measured structure footprints as a predictor, this study employs methods that run automatically and can be applied at large scale, utilizing data available in most countries.

Continuous urbanization amplifies the need for new and detailed urban data. For instance, in Mozambique, the urban population grew by more than 50 percent during the last decade, and is expected to grow roughly another 50 percent every decade for the next three decades (8). This trend is similar to many other developing countries, leading to a general urbanization of poverty(9). As in many other countries, there is no or very limited detailed data on where and how this urban growth is changing the city landscapes in-between the decennial censuses. In many cases, the extent of cities and the relative quality of neighborhoods are not recorded in systematic ways. Such shortages impede efficient social policies targeting those most in need, as well as other public policies. Standard household surveys, that are used for collection of many other vital statistics, do not cover cities in sufficient detail to provide much guidance on planning, and for most areas there is very limited administrative data available.

To address urban poverty, the Government of Mozambique is expanding an incipient urban social safety net program. In the Productive Safety Net Program (PASP), beneficiaries receive a small subsidy for supplying their labor to the production of a public good for the benefit of the local community. Hence, the program both aims to improve the livelihood of poor households as well as improve poor neighborhoods. Households are eligible to join the program based on a number of observable characteristics that result in a poverty score, also known as Proxy Mean Test (PMT) score. Households with a poverty score below a certain threshold are eligible to participate in the program. In order to facilitate a rapid scale-up this program, the government wishes to rank neighborhoods according to average poverty scores; however, they have no current data to base such a priority on.

Data and Methods

To estimate the neighborhood poverty scores, two different methods are tested. The first method directly combines households' poverty scores and the images of households' location using a Convolutional Neural Network (CNN). The second method extract information about each location from GIS data and from a CNN object detection detecting density and quality of structures, which is combined in a Random Forest model to predict poverty scores.

Data

Unit of analysis. All predictions and most of the analysis are based on a grid of 115 m x 115 m cells. The image size is 400 x 400 pixels, which measured in meters at surface level varies with a negligible amount due to the curvature of the earth. After filtering out areas with no residential buildings, the sample of interest has 57,540 cells covering five cities in Mozambique.

Poverty score (PMT) from survey data. The latest nationally representative household survey(10) as well as observations from the first round of interviews for the social protection program(11) are used to capture households' poverty scores. In both sources, the poverty score is calculated from survey questions on ownership of durable assets and other household characteristics, and both data sources include households' locations measured by latitude and longitude.

Geospatial data. The following spatial data are utilized: a) *Road data from Open Street Map (OSM)*, a volunteer-driven platform for creation of open access map data (12). This data includes the total length of - and distance to - primary, residential and other roads at the cell level. b) *Distance to city center*. To capture remoteness, the distance from the center point of each cell to the city center (as birds fly) was calculated using GIS software. c) *Estimates of structure density and quality (13).* Based on tagged images of building structures, a Convolutional Neural Network produced estimates of structure density and quality of buildings for each cell across cities in Mozambique. See SI Appendix 2 for more details.

Images. Satellite images of each cell were downloaded using Google Maps' API. Each image covers one *unit of analysis* aligning estimation and predictions into to the cell grid.

Methods

To estimate the neighborhood poverty score, two different methods are tested. The first method directly combines households' poverty scores and images of all households' locations using a CNN regression model (Fig 1). The CNN uses an architecture pre-trained to classify images in the ImageNet dataset(14, 15) and is adjusted to the images centered on the household locations. See further details in SI Appendix 2. Training of the CNN is not yet standardized, but programs facilitating easier access and applications are developing quickly, making the approach more and more accessible. The training and validation (10 percent of cells excluded sample) utilize all urban households from the two surveys (15,033 households in total) and make predictions for the 57,540 cells covering five cities of interest.

The second method predicts poverty scores from the available geospatial data; structure density and quality, the distance to city center, and length of/distance to roads using a Random Forest model (Fig 2). Capturing that not only the immediate surroundings can be important for the poverty score, spatial lags of all variables are included with weight given up to 500 meters from each neighborhood center point.

The first method is the least data demanding method, as it only relies on household survey data with information of households' poverty status and its GPS location. Such data is available in a large majority of developing countries. However, in most cases the approach requires datasets with many observations, in excess of the number of households often found in standard surveys. Further, the method, as applied here, could also be challenged in that two similar in appearance locations could have different associations to poverty. Imagine two similar locations, one close to downtown, another far from the city center. Such two locations are likely to have different associations to poverty, which is unknown to the CNN. The second method utilize data from several sources, and is more demanding on required data.

Both methods predict the cell poverty score. For both methods, the same 10 percent of the cells are excluded from modelling and used to evaluate out-of-sample prediction accuracy. The excluded sample is stratified over cities, with 10 percent randomly selected within each city.

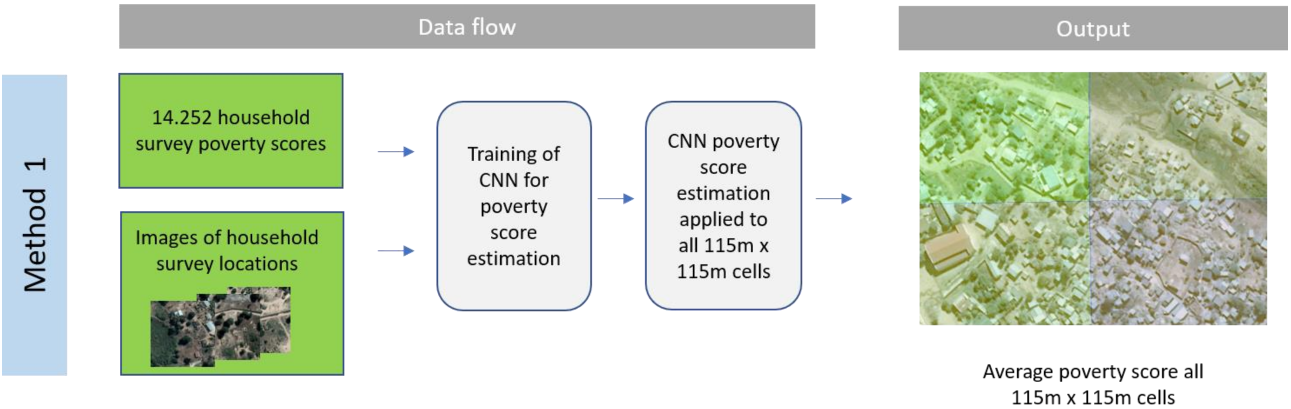


Fig 1: Figure shows the data flow of method 1. The target variable for training of the CNN is the household poverty score. See further details in S1 Appendix 2

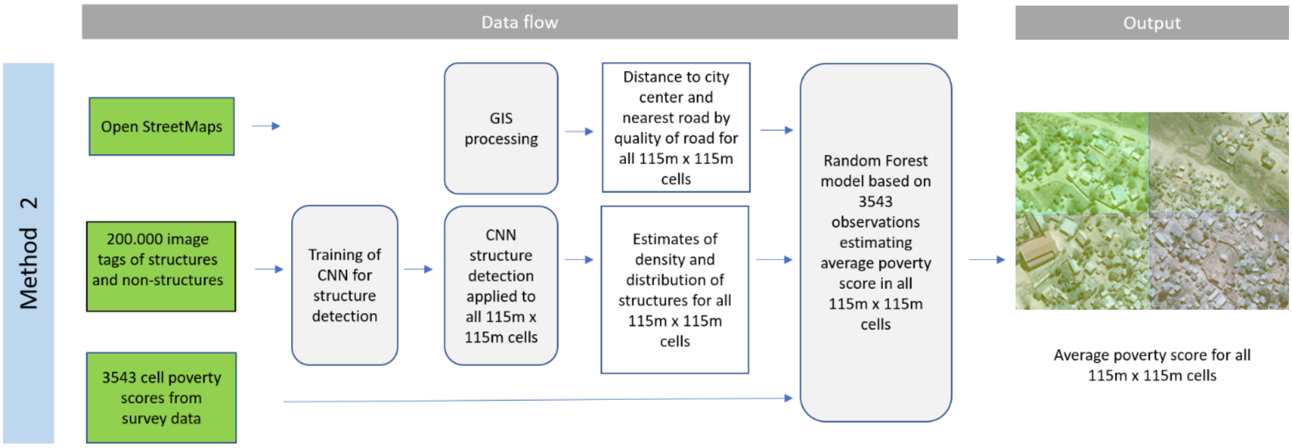


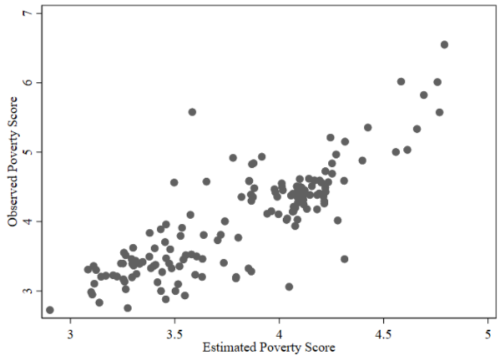
Fig 2: Figure shows the data flow of method 2. The target variable for training of the CNN is the presence and quality of structures. See further details in S1 Appendix 2.

Results

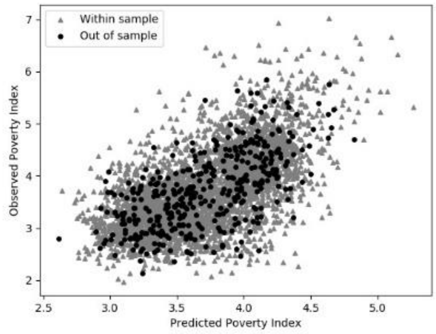
Using the direct approach and predicting poverty score only based on household survey poverty scores and images of the neighborhood in which the households are located (Fig 1), does provide insights into spatial distribution of the poverty score. The out-of-sample Spearman rank coefficient and r-square are 0.62 and 0.36, respectively (Table 1). Though not a great prediction model, this can be seen as a breakthrough for the approach, especially considering that several challenges are hampering the application. To the authors' knowledge, such direct estimation of poverty, based on images and survey data only, has only been attempted by one previous study (6). Their attempt at estimating poverty rates at sub municipal level was unsuccessful and they concluded that *"the algorithm did not learn a meaningful representation of poverty"*. Further, one should expect even better results if raw data was better, for example multi-band satellite images in higher resolution than used here. In this application, additional noise is expected as there is a time gap between observing the poverty score and the neighborhood image. Some inaccuracies were also detected in the households' location data, which likely led to additional noise in the model. Finally, image recognition relies heavily on large datasets, and this application succeeded with 14,252 training images.

Poverty score predictions based on geospatial data (Fig 2) are notably more accurate than those based on image recognition only (out-of-sample r square of 0.58 compared to 0.38) (Table 1). The model predictions reveal that the city variable for Maputo is the most important variable, next is distance to city center, and the share of high-quality roofs. In general, the estimates of structure density and quality of structures (and their spatial lags capturing the wider area) are prominent among the geospatial variables most important for poverty score predictions, while the road data is less prominent (SI Appendix 1, Fig A1-1). The geospatial model (Fig 2) highlights that information unavailable to the CNN model (Fig 1), like city and distance to city center, are important for predictions of poverty scores.

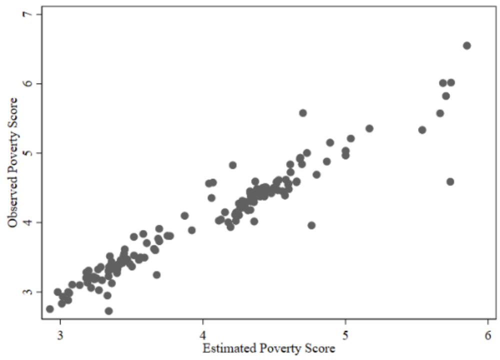
Operationally, the PASP program is implemented through *Bairro*s, a low-level administrative unit within cities in Mozambique. In fact, the original request from the Government was on assistance in ranking *Bairros* according to the poverty score. A *Bairro* ranking can be obtained by averaging cell estimates for within each *Bairro*. Evaluating model accuracy at *Bairro* level shows a remarkable out-of-sample r-square of 0.66 for method 2 (Table 1). The CNN poverty score predictions (Method 1), at this aggregated level, also have a remarkable r square of 0.58 (Table 1). Fig 3 illustrates the correlation between predicted and observed poverty scores at cell and Bairro level, including Method 2s better performance. Figure A1_3 in SI Appendix 1 exemplifies the different outputs of a cell-level and Bairro-level approach.



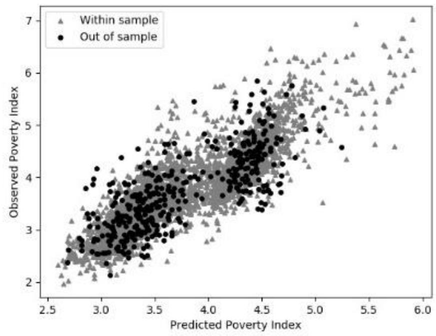
Method 1, Bairros



Method 1, Cells



Method 2, Bairros



Method 2, Cells

Fig 3 shows predicted and observed poverty score in Bairros and cells from method 1 and method 2

		R square		Spearman rank correlation	
Data	Method	Within sample	Out of sample	Within sample	Out of sample
Cells (115x115m)					
Survey, images	CNN	0.361	0.384	0.597	0.619
Survey, geospatial	Random Forest	0.750	0.576	0.886	0.757
Aggregation of cell results to Bairro level					
Survey, images	CNN	0.691	0.576	0.797	0.763
Survey, geospatial	Random Forest	0.918	0.661	0.956	0.834

Table 1 shows accuracy of predictions from three main models as defined in Fig 2. HH Survey is the household poverty surveys explained in section Data and Methods. R square is explained variance and is based on the 115m x 115m neighborhood estimates irrespective of training set. The Random Forest model is based on 150 trees and have an out-of-bag r square of 0.58.

Illustrating how this approach can provide support for the Government of Mozambique, Fig 4 shows predicted poverty scores at the cell level for the city of Maputo. The highlighted extract shows an overlay of predictions in Google Earth, illustrating the differences between relatively poor and affluent areas. It's notable that the method is capable of categorizing cells by average poverty score, even when in close spatial proximity. Maps like Fig 4 for all five cities provide easy and efficient operational support for the Government of Mozambique as they expand the urban social safety net program (PASP).

Fig 4: Left image shows poverty scores in the city Maputo based on predictions from Model 3 colored from red, poor, to blue, rich. Right image shows enlargement of section of city, illustrating the poverty score predictions for cells (115x115 meters).

Of general interest is the degree to which models are transferable, i.e. whether a model trained on one city can be applied in a different city. If not, the model would not be applicable to cities not included in the household surveys. Testing this, by excluding a full city from the training data and predicting the poverty score for the excluded city, utilizing Method 2, reveals that the models are not transferable. Predictions for the excluded city are poor. This is consistent with the city variables being prominent among most important variables (SI Appendix 1 Fig A1-1), and indicates that each city has different prediction models. For utilization on the urban social program in Mozambique, this is not a problem as the household survey covers all major cities and the model accuracy is similar in all cities, but it does reveal a current limitation of the approach. The limitation might be an artefact of having insufficient training data and is therefore surmountable.

Future Use

The core results are both a significant step forward in application of image recognition in urban settings and in use of machine learning methods, showing that they can provide key support to the Government of Mozambique and their poverty reducing operations. This is despite several data challenges, with imperfect alignment of data across time and poor quality of household tagging, leading to noise.

The frequency and quality of satellite images are increasing rapidly, leading to a large potential for continuous update of the poverty scores maps. Applying the same model based on new images, implicitly assumes that the same model is applicable over time. Though critical to the possibility of tracking SDGs and other data relevant for social targeting and planning, with such methods, this assumption has never, to the authors' knowledge, been tested empirically in the context of poverty predictions using geospatial data or image recognition. There is some experience on model stability across time from prediction of poverty based on household characteristics across surveys. Such models rely on richer data from each household including asset ownership, location, family composition and other characteristics. The experiences are mixed with some successful applications as well as failures, with little clarity on why some succeed while others fail (16, 17). Unlike survey data that can be susceptible to systematic variation in reporting over time due to variations in design and implementation(18), satellite data is generally less susceptible to such biases, though cloud cover and seasonality are potential challenges.

References

1. Talip Kilic US, Hiroki Uematsu, Nobuo Yoshida (2017) Costing Household Surveys for Monitoring Progress Toward Ending Extreme Poverty and Boosting Shared Prosperity. *Policy Research Working Paper 7951*.
2. Jean N*, et al.* (2016) Combining satellite imagery and machine learning to predict poverty. *Science* 353(6301):790-794.
3. Engstrom R, Hersh JS, & Newhouse DL (2017) Poverty from space : using high-resolution satellite imagery for estimating economic well-being.
4. Pandey SM, Agarwal T, & Krishnan NC (2018) Multi-Task Deep Learning for Predicting Poverty from Satellite Images. in *The Thirtieth AAAI Conference on Innovative Applications of Artificial Intelligence*.
5. Abelson B, Varshney KR, & Sun J (2014) Targeting direct cash transfers to the extremely poor. in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining* (ACM, New York, New York, USA), pp 1563-1572.
6. Orbital Insight (2017) World Bank: Poverty Mapping in Mexico.
7. Watmough GR*, et al.* (2019) Socioecologically informed use of remote sensing data to predict rural household poverty. *Proceedings of the National Academy of Sciences*:201812969.
8. United Nations DoEaSAPD (2018) World Urbanization Prospects,: The 2018 Revision, Online Edition.
9. Ravallion M, Chen S, & Sangraula P (2007) New Evidence on the Urbanization of Global Poverty. *Population and Development Review* 33(4):667-701.
10. ESTATISTICA IND (2014-15) Inquérito sobre Orcamento Familiar (IOF).
11. Mozambique Go (2016) Productive Social Action Programme—PASP.
12. OpenStreetMap contributors (Planet dump retrieved from https://planet.osm.org December 2018. ("\url{ https://www.openstreetmap.org }").
13. Fisker P, Malmgren-Hansen D, & Sohnesen TP (2019) TASIM: Tagged Structures in Mozambique. (Mendeley Data).
14. Russakovsky O*, et al.* (2014) ImageNet Large Scale Visual Recognition Challenge. *eprint arXiv:1409.0575*:arXiv:1409.0575.
15. Chollet F (2016) Xception: Deep Learning with Depthwise Separable Convolutions. *eprint arXiv:1610.02357*:arXiv:1610.02357.
16. Dang H-A, Jolliffe D, & Carletto C (DATA GAPS, DATA INCOMPARABILITY, AND DATA IMPUTATION: A REVIEW OF POVERTY MEASUREMENT METHODS FOR DATA-SCARCE ENVIRONMENTS. *Journal of Economic Surveys* 0(0).
17. Sohnesen TP & Stender N (2017) Is Random Forest a Superior Methodology for Predicting Poverty? An Empirical Assessment. *Poverty & Public Policy* 9(1):118-133.
18. Kilic T & Sohnesen TP (Same Question But Different Answer: Experimental Evidence on Questionnaire Design's Impact on Poverty Measured by Proxies. *Review of Income and Wealth* 0(0).
19. Kingma DP & Ba J (2014) Adam: A Method for Stochastic Optimization. *arXiv e-prints.*
20. McBride L & Nichols A (2016) Retooling Poverty Targeting Using Out-of-Sample Validation and Machine Learning. *The World Bank Economic Review* 32(3):531-550.

SI Appendix 1 Additional tables and figures

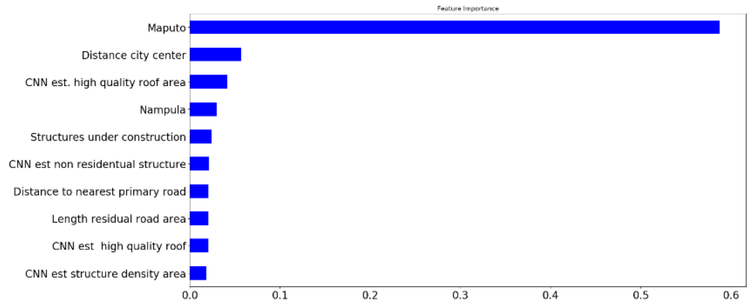


Fig A1_1. Importance scores for method 2

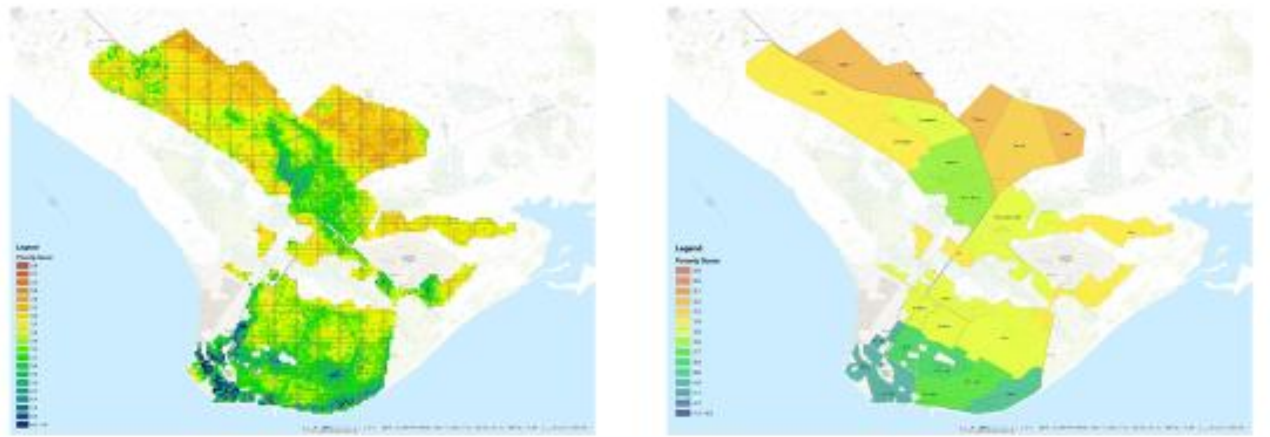


Fig A1_3: Comparison of cell-level and bairro-level outputs using the city of Beira as example

SI Appendix 2 Details on method and data

The paper utilizes two different CNN estimation models, and a Random Forest for the following purposes:

1. CNN estimation of structure density and quality.
2. CNN estimation of poverty scores.
3. Random Forest estimation of poverty scores.

The three models are described in more detail below.

CNN estimation of structures and their quality and use

Data

The training data for the CNN prediction model consist of 107.507 structures tagged into seven categories (house colored roof, house grey roof, house palm roof, non-residential structure, small non-residential structure, structure under construction, apartment). The data cover the entire city of Tete and some sections of the cities Maputo, Beira, Nampula and Quelimane. Each structure's location and type were tagged by

Cloud Factory in Google Earth in September 2016. To train an algorithm to detect structures you need both images of structures and non-structures. Based on all the identified structures in Tete, an additional 100.000 center points were identified with at least 20 meters to the nearest structure, completing the sample with locations with structures and no-structures. This is possible as all structures in Tète had been tagged.

Subsequently, an image of the location for each structure and no-structure's location was downloaded using Google Maps API in November, 2018. Images for training and predictions are from Google Maps and were accessed for free for non-commercial use, according to Google's ”Fair use” agreement. The time gap was due to a gap in funding. The tagged structure data is (*Tagged Structures in Mozambique*(TASIM)) available online(13).

CNN model

The data is utilized for two separate CNN models; first a CNN-detector estimating if a structure is present or not, and secondly, if a structure is present, a CNN classifier predicting the type of structure (see Figure A2-1). The data are split into two CNN algorithms, as it allows a more optimal use of data than a combined model.

The CNN detector is trained on the center 50x50 pixels in the approximate 200.000 images with structures and non-structures.

The CNN-detector covering 50x50 pixels is applied on the 400x400 cell images in a sliding window approach, with each window being classified as structure or no-structure. The sliding window has 75% overlap between windows, leading to 900 windows per 400x400 pixel cell. An attempt to train the CNN-detector to only detect a structure if the window is perfectly aligned over the structure, i.e. giving one detection per structure, was unsuccessful. The lack of success was likely because the TASIM geo-coordinate tags were too imprecise. Hence, the CNN detector does not provide estimates of number of observed structures as the original data, but rather the share of area covered by structures. This is referred to as structure density.

The second algorithm is a CNN-classifier that classifies windows of 50x50 pixels in 4 categories: Grey roof, painted roof, non-residential, under construction. Two categories - “Palm roof” and “Apartment” are excluded. The latter was excluded due to too few observations (192), while initial experiments with the category “Palm roof” was unsuccessful, most likely due to too many incorrect tags. To keep a balanced sample across categories, only 10 percent of the “grey roof” category was included, as this category was overrepresented (92% of the 107 507 houses). For the other categories, all available observations are used; 1267 for non-residential, 6027 for under-construction and 3648 for painted-roof.

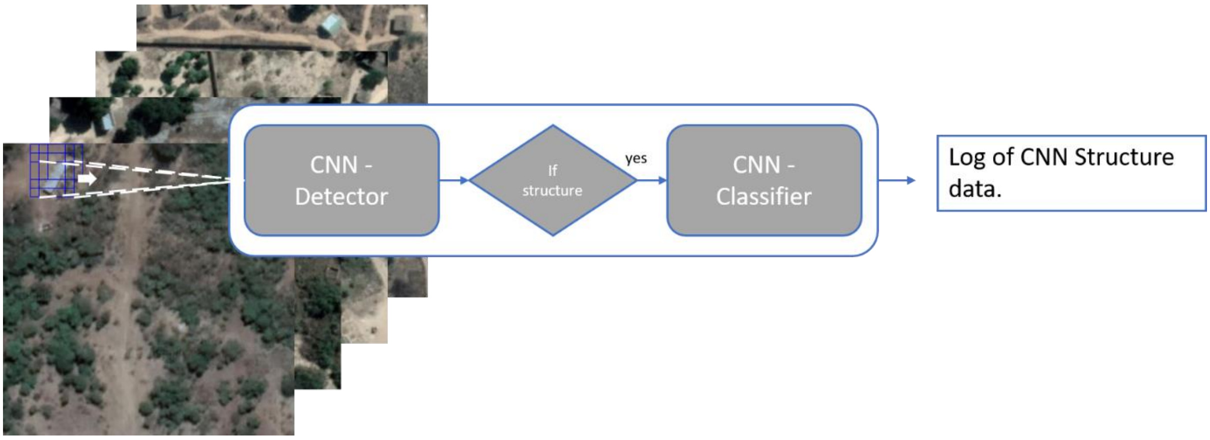


Figure A2- 1 -Results of applying the CNN model over a cell image. First the CNN Detector predicts if partial image has a structure or not. While the model is not 100% correct is has a high accuracy in detecting the structures marked by blue squares. Secondly, for each blue square the CNN classifier assigns a probability of structure type (Grey roof, painted roof, non-residential, under construction).

Both the CNN-detector and CNN-classifier were trained with the RMSprop weight update scheme for 80 and 200 epochs respectively. For the CNN-detector data augmentation was applied with 10% random shift in width or height, random rotation within 15 degree and a horizontal flip. The learning rate was initialized at 0.001 and reduced to 0.0005 after 10 epochs and to 0.0003 after 60 epochs. Data augmentation did not improve the CNN-classifier training and were therefore not used but a learning rate decay scheme like the one for the CNN-detector was used. The reduction was from 0.001 to 0.0005 happened after 75 epochs and from 0.0005 to 0.0003 after 100 epochs.

For each cell of 400x400 pixels the following parameters are derived:

1. Constructions density. The number of windows with detected structure divided by the total amount of windows per cell.
2. Model Confidence in Structure detection. The sum of the probabilities of all windows with higher than 50% probability of a structure.
3. Ratio of structures in category:
    a. Grey roof.
    b. Painted roof
    c. Non-residential
    d. Under construction

In the 10 percent of observations excluded for evaluation, the CNN detector has an accuracy of 97% in detecting structures, while the CNN classifier has an out-of-sample test accuracy of 69%. The confusion matrix in Table A2-1 shows the accuracy for each class, while Figure A2-2 shows that there is a high correlation between predicted structure density and the number of tagged structures within each cell. Note that the latter are two different variables and not a prediction of the same variable.

The CNN detector and classifier is implemented in Python using the *Keras* library.

	Grey roof	Painted roof	Non-residential	Under constr.
Grey roof	79%	3%	3%	15%
Painted roof	11%	73%	3%	13%
Non-res.	22%	24%	50%	4%
Under constr.	23%	7%	1%	69%

Table A2-1 – Confusion matrix on the performance on the CNN-classifier. Percentage correctly classified.

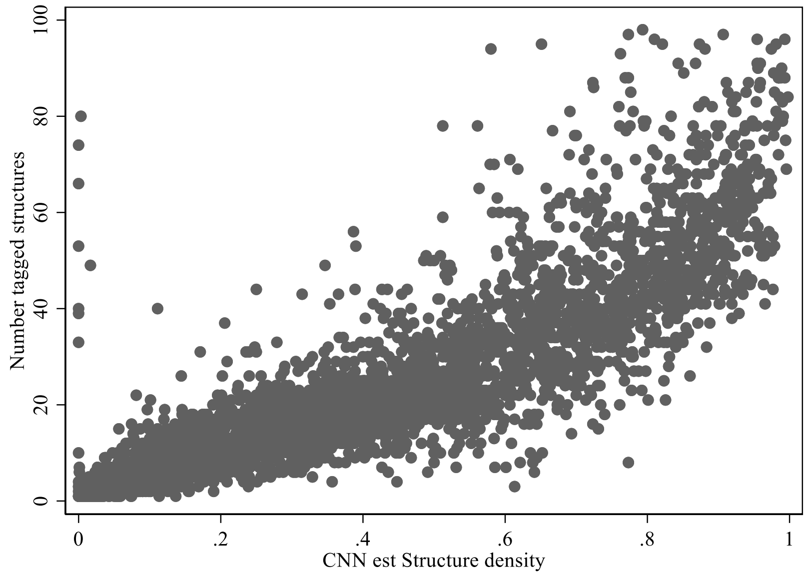


Figure A2-2 - Plot of structure density parameter from CNN algorithm vs. the tagged structures in each cell in Tete.

CNN estimation of poverty scores

Data

The training data for the CNN model of poverty scores is based on the IOF and PASP surveys (10, 11). Both surveys have data on each households poverty score as well as their geo-coordinates and combined they have 14.252 urban household observations. For the training, all urban households are included, including observations outside the five cities of interest.

An image of the location for each household was downloaded using Google Maps API in November, 2018. Images for training and predictions are from Google Maps and were accessed for free for non-commercial use, according to Google’s ”Fair use” agreement.

CNN model

Estimates of poverty scores directly from images is based on a CNN designed for regression (referred to here as the CNN-regressor). Training a CNN from scratch with 14.252 images gave unsatisfactory error rates. Evaluation led to the conclusion that lack of data was the reason.

To overcome the small sample, the following steps were taken:

1. Transfer Learning: The CNN architecture Xception(15) which is pre-trained to classify images in the ImageNet (14) dataset was used as the basis for training.
2. Cropping: Five crops of each image were used, top-left, top-right, bottom-left, bottom-right and center.
3. Data augmentations: each image was flipped, and a random scaling of its' brightness level of +/- 15% were applied.

The CNN-regressor was trained with the ADAM weight update scheme with parameters as recommended in Kingma and Ba (19) for 20 epochs. A learning rate decay was also applied reducing the learning rate with 2% for each epoch. The CNN-regressor minimizes the Mean-Square-error of the difference between the predicted and recorded poverty scores.

Final CNN-regressor predictions in cells are based on five crops and take the mean of these five predictions as the final prediction for each cell. This is in order to cover the whole area of the 400x400 pixel cells as the pretrained CNN was designed for images of 299x299 pixels. Figure A2-3 illustrates the cropping.

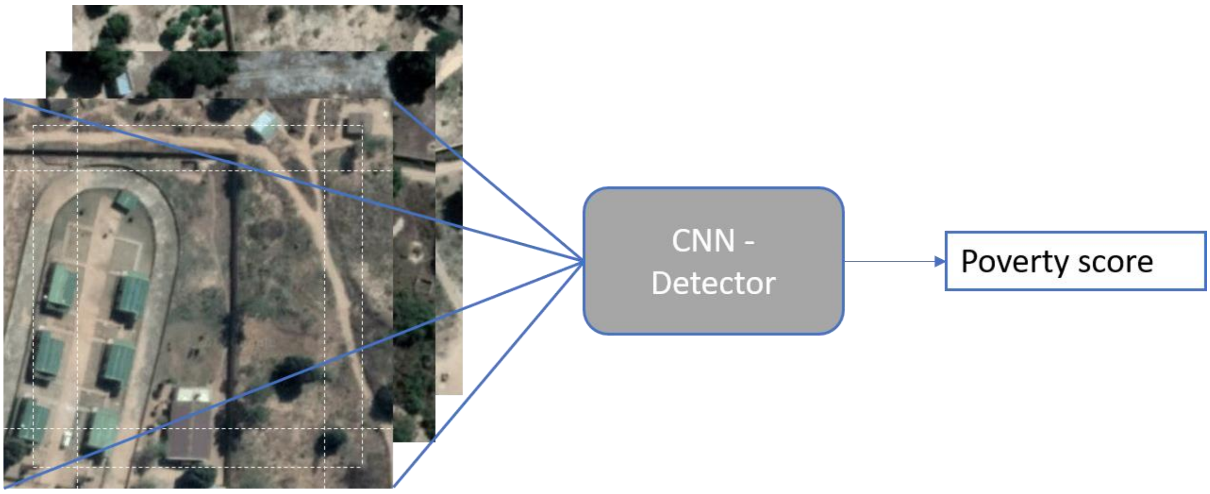


Figure A2-3 - Applying the CNN-regressor for poverty score prediction over cell level images. White dashed lines indicate the 5 crops from which an average poverty score is taken.

This yielded a Mean-Square-Error (MSE) on the training dataset of 0.02 and an out-of-sample MSE of 0.35. The test accuracy is calculated on 10% of the cells with poverty scores. Test set cells was selected in a stratified manner over the five cities of interest.

See also Figure A1-2 and Table 1 for further results on model prediction accuracy on cell level. Note the excluded sample is based on cell level, but data set vary between training and predictions.

The CNN regressor is implemented in Python using the Keras library.

Random Forest estimation of poverty scores

Data

The Random Forest model prediction models are based on the 115x115 m cells defined for the five cities covered. The average poverty score from IOF and PASP households in each cell is the target variable the models. The full sample is 3,927cells, of which 10 percent (392 observations) is excluded for evaluation, leaving 3,535 cells for training of the models. The excluded sample is stratified over the five cities.

Features, at cell level, for the model include:

- Length and distance to nearest primary, residential and other road from OSM.
- Distance to city center.
- CNN Estimates of structure density and quality and use, as described above.
- CNN estimates of poverty score, as described above.
- Spatial lag of above features with exponential distance decay parameter of 1.5 including up to 500 meters from center points of each cell.

Random Forest prediction of poverty score

Application of Random Forest is recent and still scant for poverty predictions, though evaluation of the method compared to alternatives have been found favorable (17, 20).

The predictions were implemented in the Python anaconda package using the RandomForestRegressor module from sklearn.ensemble. Mean square errors was used as decision criteria, minimum leaf size was set at 10, and predictions are based on 150 trees.