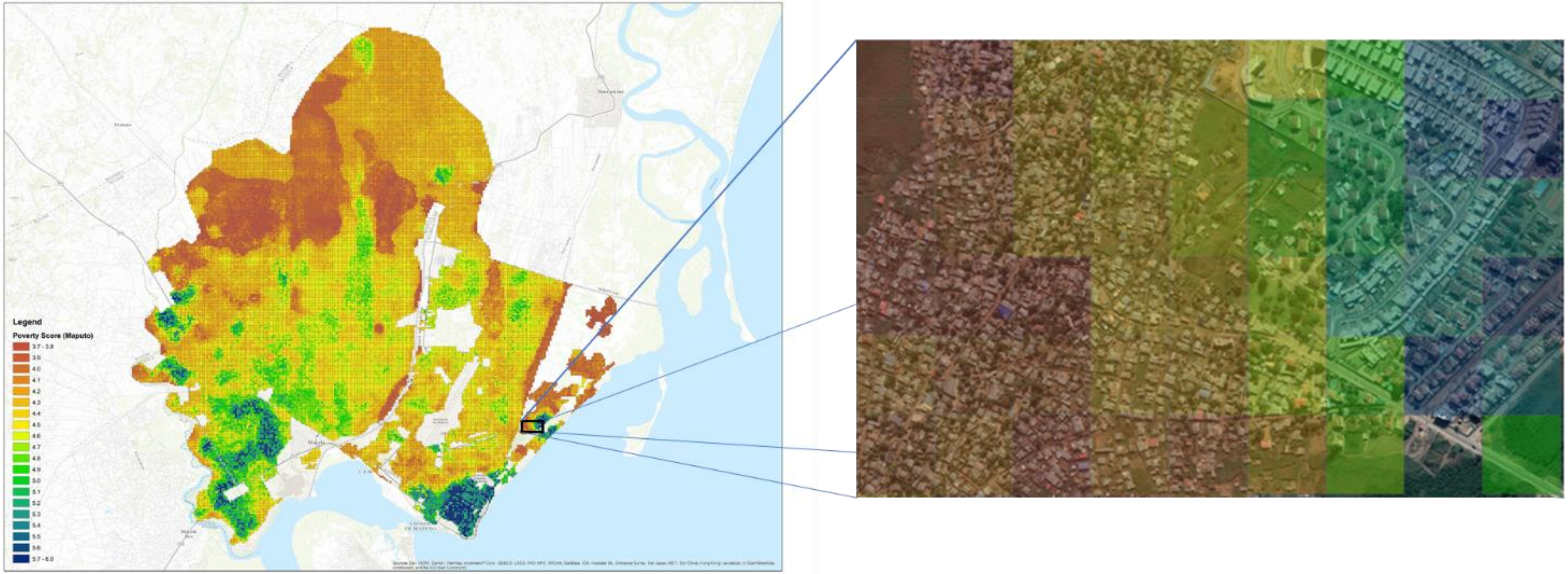
Using Satellite Data to Guide Urban Poverty Reduction

Peter Fisker (University of Copenhagen)
David Malmgren-Hansen (Technical University of Denmark)
Thomas Sohnesen (Consultant/Advisor World Bank, UNICEF, University of Copenhagen, UNU-WIDER)

The output: Average Poverty Score (PMT score) in a resolution of 115mx115m for all major cities



How did we get there?

We empirically test two different approaches:

Method 1:

- Using convolution neural networks (CNN) to estimate Poverty Scores directly from satellite images

Method 2:

- Using machine learning based Random Forest prediction model based on the following inputs:
  - Density and quality of building structures, extracted from satellites images using a Convolutional Neural Network (CNN).
  - Distance to center of city and nearest road of various quality, extracted from OpenStreetMaps

Short on Convolutional Neural Network (CNN)

- Convolutional Neural Networks (CNN) has gained popularity in recent years for a range of computer vision tasks.
- They excel in image related tasks such as classification, object detection and segmentation.
- They work by applying a set of filtering operations to an image which enhance different features of the image context. Optimal feature filters are then found by optimizing the model over a set of training data samples rather than previous approaches where the filters were manually engineered.
- Pros: Many new and interesting applications.
- Con: Needs large amount of training data.
  - Transfer learning is allowing application for smaller data sets.

Short on Random Forest

- Random forest is a widely applied prediction method.
- The random forest method is part of the machine learning literature and utilized for predictions in a wide range of research fields, including medical.
- The application of random forest is recent and still scant for poverty predictions, though evaluations of the method compared to alternatives have been found favorable.
- Predictions rely on a large number of prediction models, each model based on subsets of observations and a subset of predictors.
- Pros: better at predictions, more robust and less prone to overfitting.
- Cons: Less of a theoretical framework, not able to do hypothesis testing.

Method 1

Data flow

Output

Method 1

14.252 household survey poverty scores

Images of household survey locations

Training of CNN for poverty score estimation

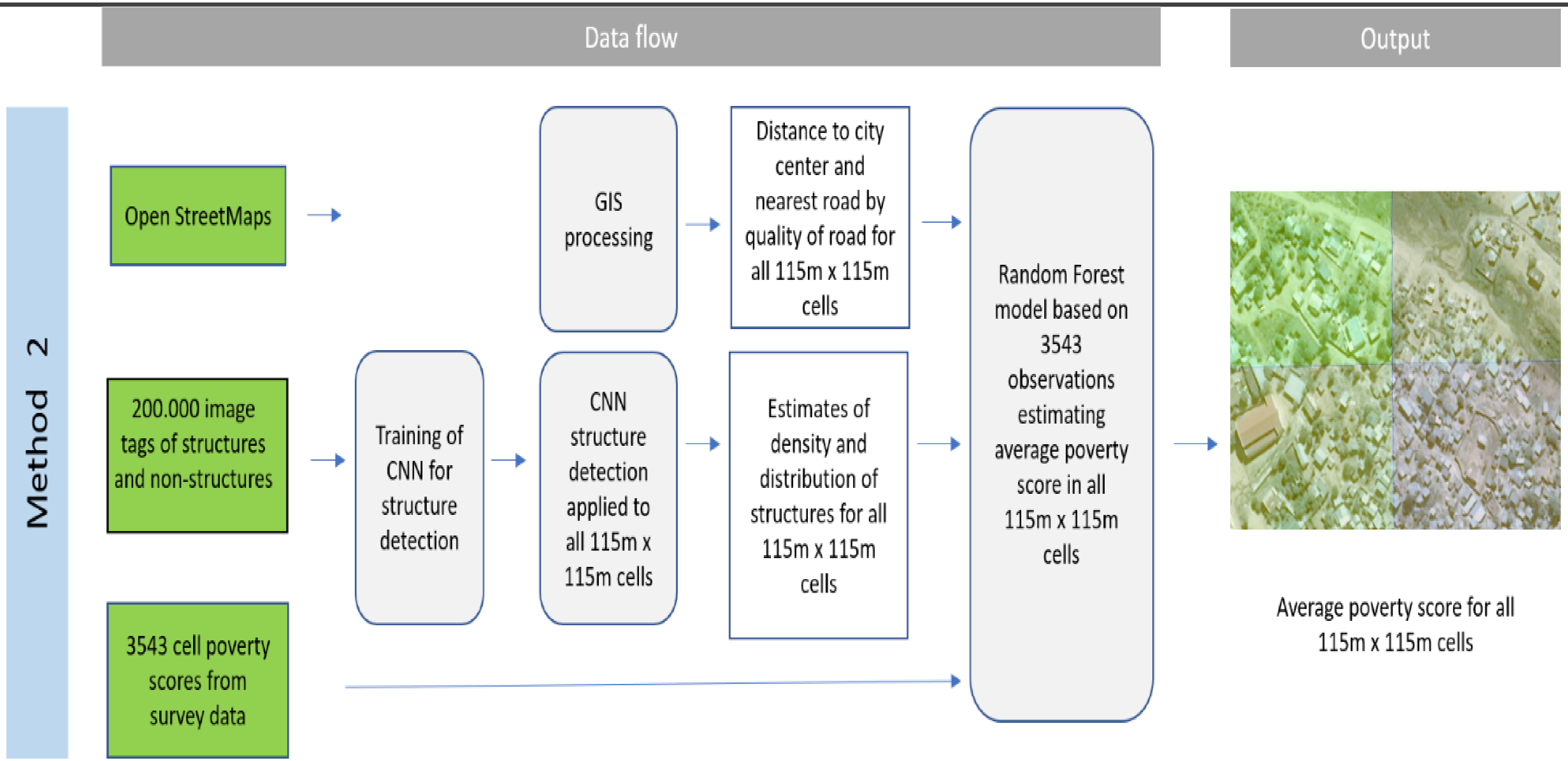
CNN poverty score estimation applied to all 115m x 115m cells

Average poverty score all 115m x 115m cells

CNN regressor estimation of Poverty Scores (PMT scores)

- Target data is PMT scores for 14.252 urban households in IOF and PASP surveys.
- Images for these 14.252 households were downloaded from Google Maps based on GPS.
- Training a CNN from scratch on these 14.252 images was unsuccessful, due to too little data.
- Using transfer learning from an existing CNN architecture, pre-trained to classify images in ImageNet, combined with cropping and data augmentations, gave notable better results.

Method 2


Data flow
Output
Method 2
Open StreetMaps
200.000 image tags of structures and non-structures
3543 cell poverty scores from survey data
Training of CNN for structure detection
GIS processing
CNN structure detection applied to all 115m x 115m cells
Distance to city center and nearest road by quality of road for all 115m x 115m cells
Estimates of density and distribution of structures for all 115m x 115m cells
Random Forest model based on 3543 observations estimating average poverty score in all 115m x 115m cells
Average poverty score for all 115m x 115m cells

Training of CNN for structure detection

Raw data:

- 107,507 tagged structures in the following categories:
    1. House with colored roof
    2. House with grey roof
    3. House with palm roof
    4. Non-residential structure
    5. Structure under construction
    6. Apartment building
- 100,000 non-structures center points - with at least 20 meters to the nearest structure.
- Images of the 207,000 locations with structures and no structures, downloaded from Google Maps API.
- In final training
    - Palm roof and Apartment had too few observations and were excluded.
    - Grey roof" category is largely overrepresented (92% of the 107 507 houses) only a subset of 10 000 images from this group was used.

Example structure detection

- By applying the CNN-detector in a sliding window approach over each of the 400x400 pixel cell images, the CNN-detector provides a count for number of windows with more than 50% probability for a structure, leading to an estimate of structure density.

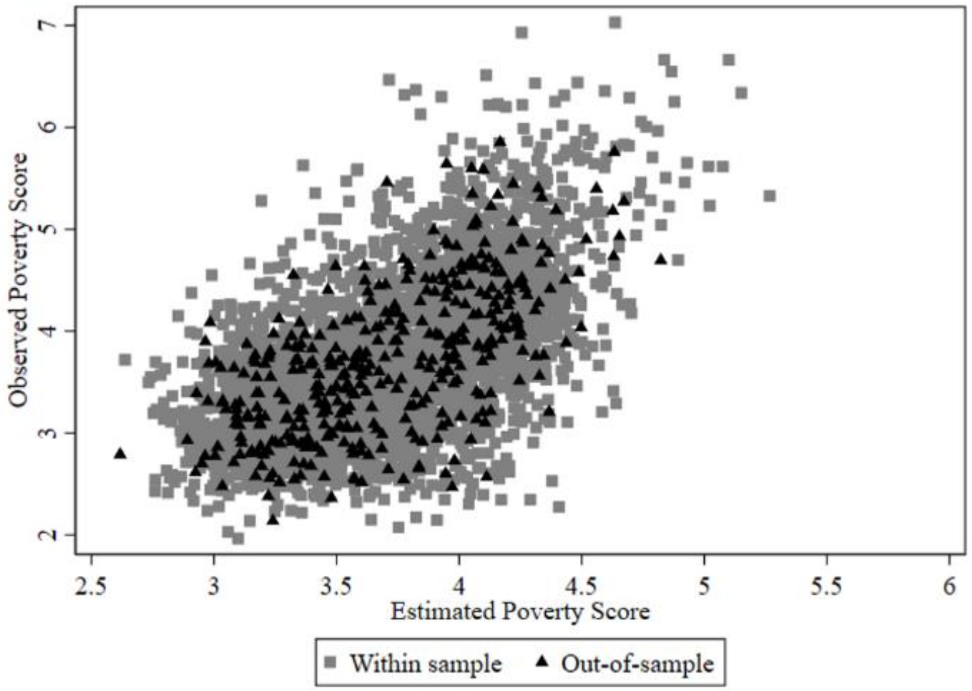
Accuracy of CNN structure estimates

- Within test set accuracy of structure detection was 97 percent
- Within test set accuracy of type of structure was 67 percent, with this distribution:

	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%

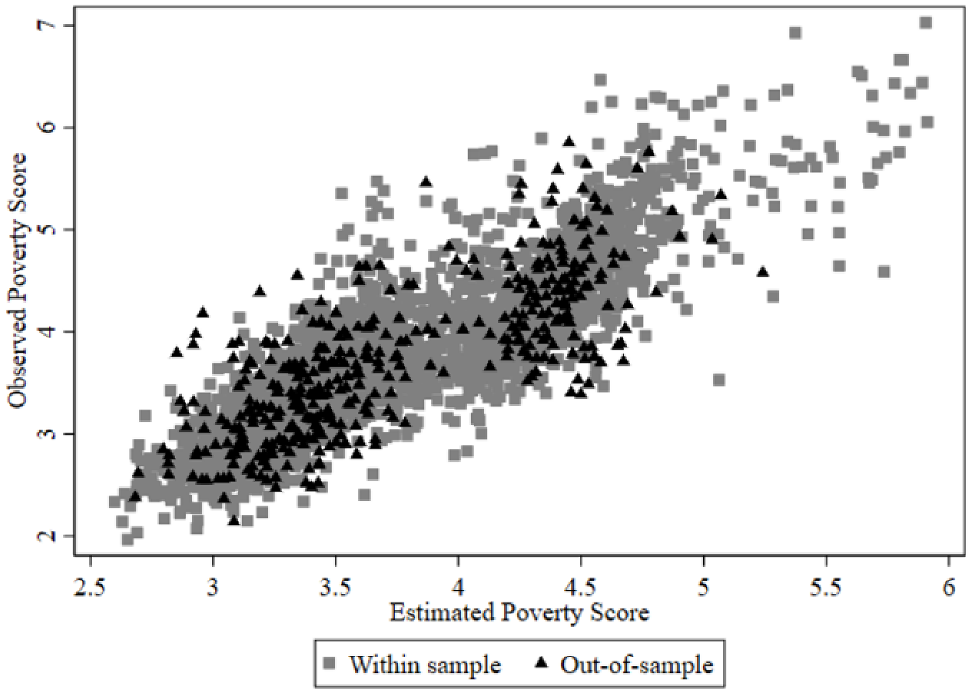
Accuracy of prediction models at cell level

Method 1 Cell level



R square		Spearman rank correlation	
Within sample	Out of sample	Within sample	Out of sample
0.36	0.38	0.60	0.62

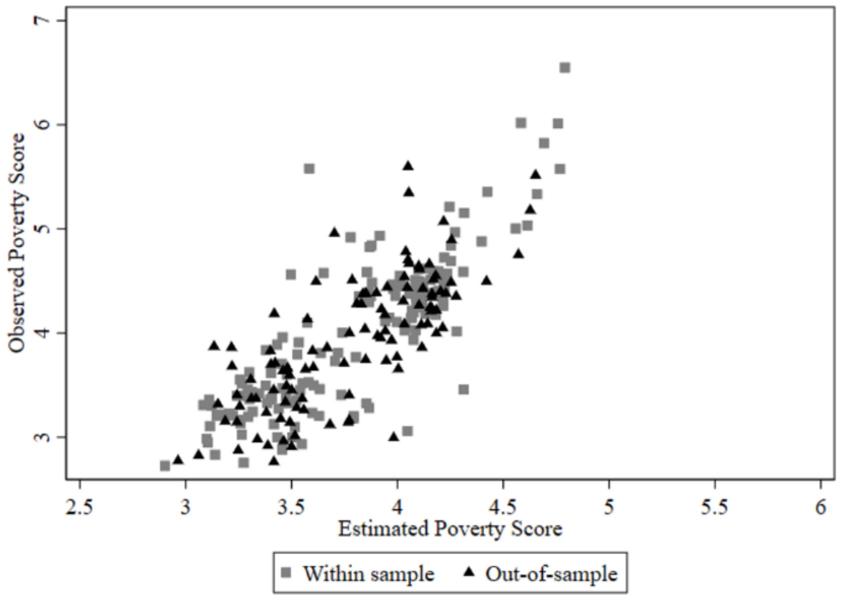
Method 2 Cell level



R square		Spearman rank correlation	
Within sample	Out of sample	Within sample	Out of sample
0.76	0.58	0.89	0.76

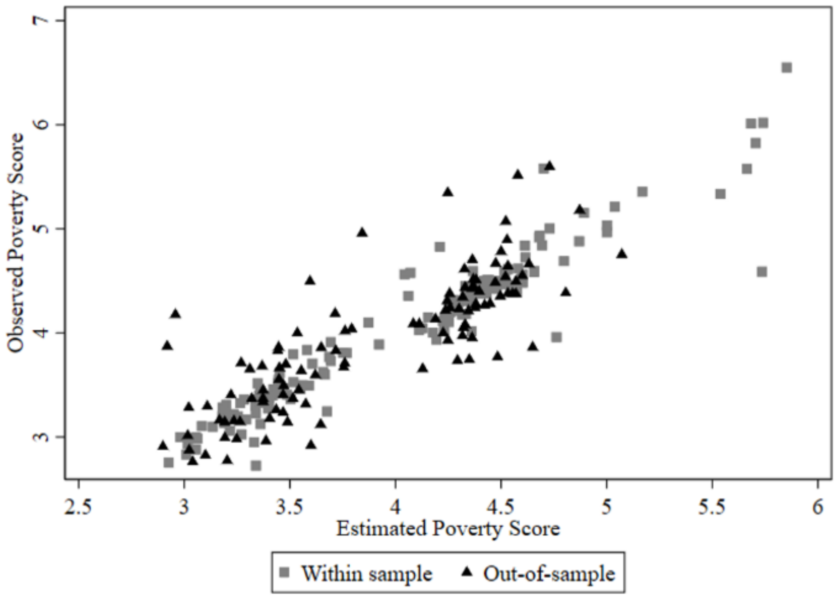
Accuracy of prediction models at Bairro level

Method 1 Bairro level



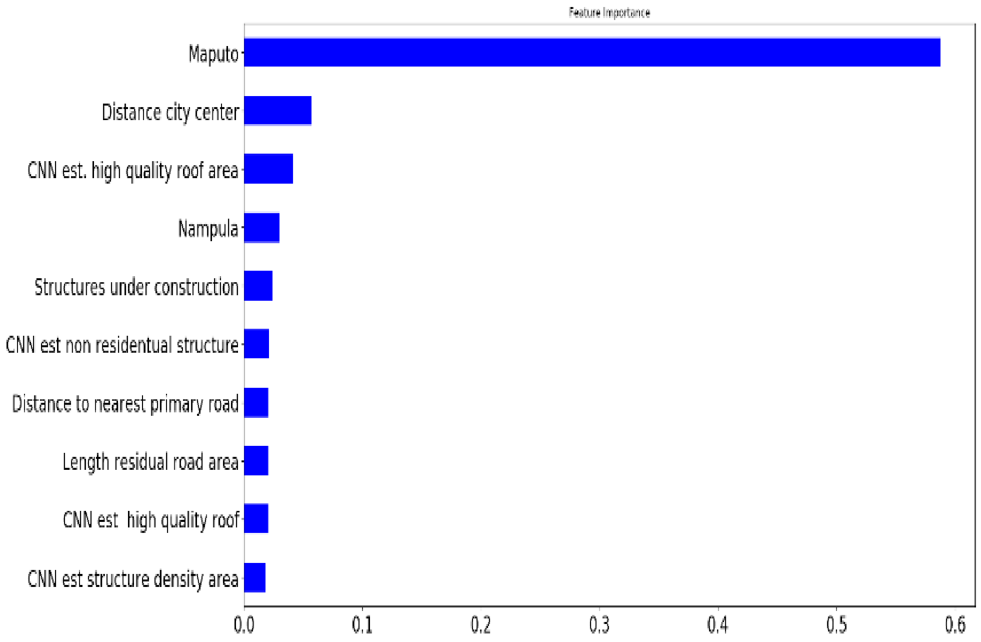
R square		Spearman rank correlation	
Within sample	Out of sample	Within sample	Out of sample
0.69	0.58	0.80	0.76

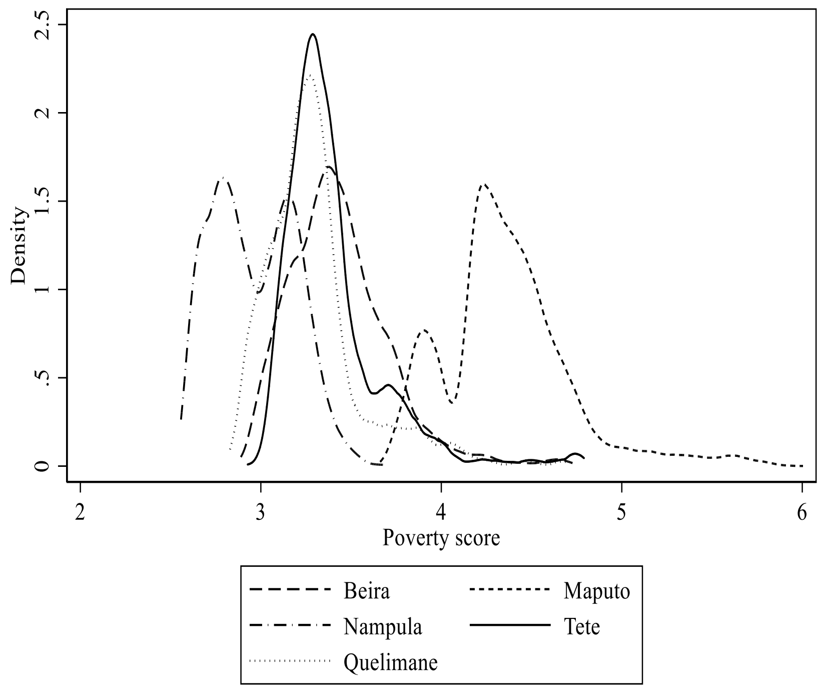
Method 2 Bairro level



R square		Spearman rank correlation	
Within sample	Out of sample	Within sample	Out of sample
0.92	0.66	0.96	0.83

Importance scores for method 2





Further results and observations

- The models are not transferable, as predicting a city excluded from training was unsuccessful.
- There is room for improvement in raw data, as we have noise due to:
  - Data is imperfectly aligned across time
  - The quality of structure tagging could be better

Moving forward we would like to:

- Do further testing of accuracy and robustness of results, and develop standard errors, utilizing known data.
- Gain more experience with sample size and city variation, exploring
  - Other countries with geo referenced data
  - Application to multicounty settings
  - Model transferability across cities and countries
- Test application of method on other areas: population, health, education, transport, infrastructure.
- Test potential use of CNN structure estimator for impact assessments, for instance road construction and disaster.

Thank you

How expensive is this?

- The tagging of images was USD 4.800.
  - Based on gained experience, we believe it can be done of higher quality at lower costs.
- Updating predictions for Mozambique based on new Google Map images cost very little.
  - Issue: Google maps images are not dated, and the date of current images is unknown, as is the date of next update.
  - Cost of purchasing images vary notably and must be assessed on a case by case basis.