MEASURING CHANGES IN POVERTY WITH DEEP LEARNING AND SATELLITE IMAGES

Lukas Kondmann, Xiao Xiang Zhu *
Technical University of Munich (TUM), 80333 Munich, Germany
German Aerospace Center (DLR), 82234 Wessling, Germany
{lukas.kondmann,xiaoxiang.zhu}@dlr.de

ABSTRACT

Timely and accurate data about local wealth is essential in the global fight against poverty to allocate resources efficiently. In many countries, such data is unavailable or outdated. To close this information gap, predicting local poverty with daytime satellite imagery and machine learning has shown promising results. While this approach excels at predicting relative *levels* of local poverty, it is, however, unclear if it can also be used to monitor *changes* in poverty over time. In this paper, we test this property and find that the approach struggles to capture changes in local development in Rwanda from 2005 to 2015. The strategy predicts no change in average wealth levels even though Rwanda has seen large reductions in local poverty in this period. To be able to detect the impact of interventions on poverty over time, further refinement of this method might be necessary.

1 INTRODUCTION

Ending poverty is the first of the 17 Sustainable Development Goals (SDGs) by the United Nations. In practice, however, policy-makers in low and middle-income countries (LMIC) tasked with alleviating poverty, often lack sufficient data about local development to use their resources effectively (Jerven, 2013). The main reason for this absence of reliable poverty measures is that the data collection of local poverty statistics is costly and strenuous (Blumenstock, 2018).

To close this information gap, Jean et al. (2016) present a novel way to predict local poverty based on daytime satellite images and deep learning. Their algorithm learns abstract features from predicting the nightlight activity of a daytime image since nightlights are a good proxy for human development (Weidmann & Schutte, 2017). This knowledge is then transferred successfully to predicting poverty from daytime images in Rwanda, Uganda, Nigeria, Tanzania, and Malawi.

The robustness of their transfer learning strategy has also been verified in other countries (Head et al., 2017) and with lower-resolution images (Perez et al., 2017). However, transfer learning has not been widely applied in practice for poverty mapping because it is unclear if this strategy can also capture *changes* over time and not only relative *levels* of local wealth. Measuring changes is, however, crucial for policy-makers that aim to evaluate the impact of large-scale poverty alleviation interventions such as conditional cash transfer programs:

> 'A lot of this work is using satellite data [...] to try to predict poverty at a granular level for entire countries or continents. At the moment such data seems useful for descriptive work, but it is unclear whether accuracy is enough to measure changes well over time. [...] There may not be enough signal to be able to detect the impact of interventions.'
> — David McKenzie, Lead Economist World Bank Development Group[1]

This paper aims to eliminate the uncertainty about the ability of transfer learning to measure changes in economic development. While we rely on methodology from remote sensing and computer vision,

*Corresponding Author

[1] https://blogs.worldbank.org/impactevaluations/how-can-machine-learning-and-artificial-intelligence-be-used-development-interventions-and-impact

our main contribution is to analyze the usability of the obtained poverty predictions in studies of local human development. We address the question if the transfer learning strategy can measure absolute changes in average local wealth in Rwanda based on Demographic and Health Surveys (DHS) and Landsat 7 satellite imagery.

We show that these concerns are justified because the transfer learning strategy is not able to detect significant changes in average local wealth in Rwanda from 2005 to 2015. When analyzing the predictions on a province level, the transfer learning strategy predicts a significant decrease in average local wealth for 2 out of 5 provinces between 2005 and 2015.

The fact that it does not predict *any* increases in average local wealth on the country level is concerning given the clear trend in economic development as one of Africa's fastest-growing economies. Further, the transfer learning approach has an established ability to measure levels of poverty in Rwanda well (Jean et al., 2016; Perez et al., 2017). Hence, the predictions based on transfer learning are inconsistent with province and national level economic trends in Rwanda from 2005 to 2015. This result indicates that bringing the transfer learning method to applications in policy evaluation is, at least in its current form, premature.

2 Data

Data on nightlight activity is taken from the National Oceanic and Atmospheric Administration (NOAA) annual time series on worldwide nighttime luminosity (DSMP-OLS). DSMP-OLS covers nightlight activity from 1992 to 2013 on 30 arc second grids which is equivalent to roughly 1km $\times$ 1km. Every cell has an average nightlight intensity value on a scale from 0 to a maximum of 63. We use nightlight activity from 2013 as a proxy for 2015 which is the time of data collection of the survey data and the satellite imagery.

Jean et al. (2016) use Google Static Maps Imagery for their transfer learning strategy. This source of imagery, however, only allows one to query one recent image of a region and there are no available photos from the past. To adapt the transfer learning strategy to time-series analysis, we follow Perez et al. (2017) and use yearly composite images of Landsat 7 as input images for poverty mapping. These composites are the result of taking the median of all cloud-free pixels of the images within the year in question. We only use the RGB channels of the images since Perez et al. (2017) show that going beyond the visible spectrum for poverty mapping gives at best minor improvements.

Poverty baselines are taken from household asset wealth scores in the Demographic and Health Surveys (DHS) in Rwanda from 2005, 2010 and 2015. Households are surveyed in geographic clusters such that all households within a cluster share the same coordinates. The wealth index is the result of taking the first principal component of survey responses about asset ownership and normalizing the component within each country and year (Rutstein & Johnson, 2004). This within-country normalization implies that the absolute indices are not comparable across years. They are merely a relative ranking of households within a survey.

3 Methods

Following Jean et al. (2016) we use daytime images of 1km $\times$ 1km to predict the corresponding local nightlight intensity in three classes: Low, medium and high. After the transfer learning task, we discard the softmax layer and use the extracted features as input to a ridge regression to fit average cluster poverty levels.

As large parts of Rwanda have a nightlight intensity of 0, we sample training imagery from all over East Africa (Figure 1). In total, we rely on 65,000 images of which 40,000 are randomly sampled from Rwanda. Note that we do not restrict these images to be outside of poverty cluster locations which, if anything, should improve performance. The frequency of classes is adjusted to be similar to Perez et al. (2017) where the ratio of low to non-low is about 2:1.

As feature extractors we use ResNet 50 (He et al., 2016) pre-trained on Imagenet. We allow fine-tuning of all layers in the transfer step to adjust the feature extractor to the task of predicting human development.

Figure 1: Locations of training images
Notes: Supplementary image locations to solve the class imbalance. The point size is larger than in reality for visibility purposes. Hence points may overlap in the figure but image locations are distinct.

We train the model for 25 epochs and use ADAM as optimizer (Kingma & Ba, 2015) with a learning rate of 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$ and choose the model with the best validation accuracy as extractor.

As cluster wealth across time is not comparable and surveyed households are different, we can not directly subtract wealth levels to get growth rates. Instead, we obtain wealth predictions for the whole of Rwanda on a nightlight cell level *relative to 2005 clusters*. This allows subtracting predictions for 2010 and 2015 from 2005 levels to get a predicted change in units of 2005 household wealth. We compare the predicted change rates with change in nightlights which are known to capture economic trends well (Henderson et al., 2012). Further, we compare the predicted changes on a province level to strong poverty reduction trends in Rwanda based on income data from the National Institute of Statistics of Rwanda (NISR).

As a necessary intermediary step, we first fit the extracted features from the cluster images to poverty levels in a ridge regression in 2005. Then, we verify the ability of the trained feature extractor to predict poverty in the same locations in 2010 and 2015.

4 RESULTS

The results in Table 1 show that it is indeed possible to infer wealth levels in years outside of training. Transfer learning with ResNet 50 explains 69% of the wealth variation in 2010 and 57% in 2015. This is in the range of using mean nightlights per cluster directly as predictor for local wealth and strongly above using an off the shelf pre-trained ResNet as feature extractor.

Table 1: Training Out-Of-Year Prediction on 2005 Only

	Mean Training R^2 2005	Test R^2 2010	Test R^2 2015
Transfer Learning & ResNet50	0.47	0.69	0.57
Mean Nightlights per cluster	0.52	0.74	0.61
ResNet50	0.55	0.43	0.36

Training Scores are averaged over 10 runs of 10-fold cross validation.
Test scores are calculated on the full sample.

Hence, we proceed to predict local poverty for 2005, 2010 and 2015 across Rwanda on nightlight cell level relative to the wealth index of households surveyed in 2005. Figure 2 presents sector-level change rates from 2005 to 2015. The plot points out that there are large differences in the prediction of changes from 2005 to 2015 depending on the prediction strategy used. Nightlights, in Panel

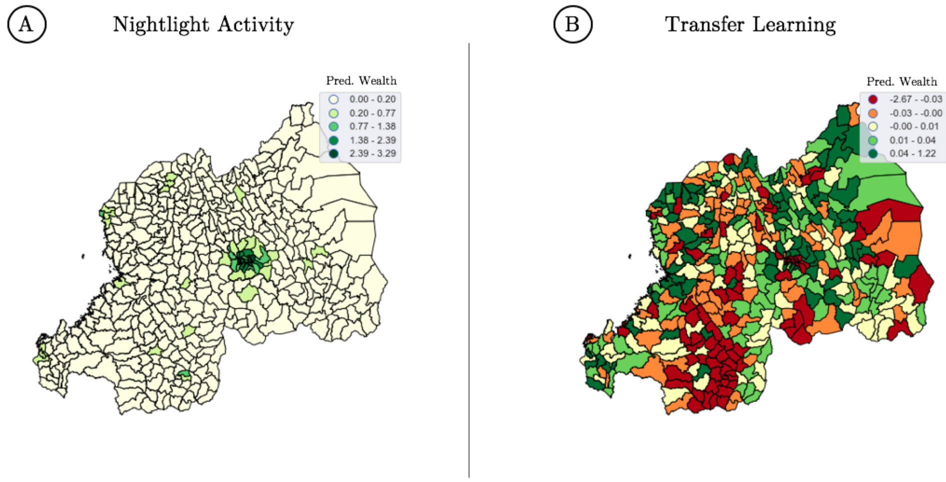

Figure 2: Predicted Absolute Change in Sector Level Wealth from 2005-15
Notes: Legends for nightlights and transfer learning are based on percentiles of the predictions and hence do not coincide perfectly.

Table 2: Predicted Average Absolute Change in Wealth Index for Rwanda

Time	Change TL	Change NTL	Poverty Reduction	GDP Growth
2005 - 2010	0.00	0.03***	0.11	0.49
2010 - 2015	0.00	0.02***	0.07	0.44
2005 - 2015	0.00	0.05***	0.18	1.15

Change different from zero for significance level: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Poverty reduction rates for 2005-2010 taken from the World Bank (2015)
and for 2010-15 from the National Institute of Statistics of Rwanda (NISR)

A, locate the largest increases around Kigali while generally, no parts of the country seem to have decreasing wealth levels.

In contrast, daytime-image based transfer learning predicts a negative trend in and around Kigali and in South-Western parts of Rwanda which is inconsistent with the nightlights prediction. Further, the predictions vary notably across neighboring sectors which creates a diverse picture of predicted changes in local wealth. For a significant part of the country, the transfer learning model predicts decreasing wealth with only some parts which are associated with positive developments. To analyze this discrepancy in more detail, I average the cell-level changes to the country level and compare the mean changes in wealth to other indicators of recent economic development in Table 2.

On average, the transfer learning approach (Column 1) predicts no increase in the average wealth score for any of the periods considered. The predicted changes are statistically indistinguishable from 0 which is tested with a one-sample t-test. On the contrary, nightlights (Column 2) predict a significant increase in all three periods. 0.03 for 2005 - 2010, 0.02 for 2010 - 2015 and hence in total 0.05 for the whole period. All of these changes are significant on the 0.1% level. The change prediction of transfer learning with daytime images is drastically at odds with nightlight activity. Further, the prediction of the transfer learning model is inconsistent with other indicators of economic development (Columns 3 & 4).

One key concern for the results in Table 2 is that the unweighted aggregation process of geographic cells might cloud the results. If the change was most notable in urban areas it could be watered down by geographically much larger rural areas. Hence, we repeat the analysis of Table 2 for the province level (Table 4) from 2005 to 2015. On average, transfer learning predicts a large but insignificant reduction in wealth in Kigali which is inconsistent with province level poverty data

Time	Change TL	Change NTL	Poverty Reduction
Kigali City	-0.04	0.83***	0.11
Eastern Province	0.01***	0.01***	0.16
Northern Province	0.02***	0.01***	0.09
Southern Province	-0.04***	0.04***	0.21
Western Province	0.02***	0.02***	0.12

Significance level: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Poverty reduction rates for 2005-2010 taken from the World Bank (2015)
and for 2010-15 from the National Institute of Statistics of Rwanda (NISR)

based on income and prediction based on nightlights. While the true change in Kigali is potentially overestimated by nightlights because of overglow the direction is consistently estimated. Hence, if anything, the transfer learning strategy would predict an even lower country level growth rate if adjusting for population in the aggregation. Further, the negative and significant change rate for the Southern Province is discouraging since nightlights and income data show a large reduction in local poverty in the time in question. Results in the Northern, and Western Province look consistent but this impression disappears when splitting up the 10 year period in 2 as in Table 2. These findings underline that the small country level growth rate is the result of noisy but not small province level growth rates that seem to be especially distorted in urban regions.

5 Discussion

We underline that the transfer learning strategy struggles to recognize change in local economic development in Rwanda. What are the reasons for this inability? One possible option is that aggregates of daytime satellite images are stable over time by design and undergo little change. It seems plausible that there is simply not enough signal of economic change in the imagery at hand. Jean et al. (2016) show that their extracted features correspond to, for example, roads or urban areas. It might be fundamentally difficult to spot changes in these variables because economic progress does not necessarily mean that there will be *more* roads. The improvement might be more subtle such as *better* roofing or road materials. This might be especially relevant for this paper because Landsat 7 images have fairly low resolution.

Nevertheless, our findings strongly indicate that the transfer learning strategy, at least as configured in this paper, can not be used for change detection in wealth in Rwanda in the analyzed period. At the moment, these conclusions are limited to the specifics of this case study in Rwanda and might fail to generalize. One could, however, hypothesize that similar results might be found elsewhere since the strong economic growth and the established ability of transfer learning to measure wealth levels in Rwanda should work in favor of detecting changes.

6 Conclusion

Our results indicate that pioneering methods that map poverty from satellite images with deep learning may struggle to capture trends in economic development over time. Careful validation of these outcomes in other countries and with other imagery is necessary to speak to the robustness of this weakness. This ability, however, is a key element that local measures of economic activity have to fulfill to help researchers and practitioners understand what interventions and policies help to tackle global poverty. More promising routes to measure local changes might involve data fusion or refining CNN architectures to predict changes in local activity directly.

REFERENCES

The World Bank. Rwanda poverty assessment, 2015.

Joshua Blumenstock. Don't forget people in the use of big data for development. *Nature*, 561: 170–172, 2018.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016.

Andrew Head, Mélanie Manguin, Nhat Tran, and Joshua Blumenstock. Can human development be measured with satellite imagery? In *Proceedings of the Ninth International Conference on Information and Communication Technologies and Development*, pp. 1–8. ACM, 2017.

J Vernon Henderson, Adam Storeygard, and David N Weil. Measuring economic growth from outer space. *American Economic Review*, 102(2):994–1028, 2012.

Neal Jean, Marshall Burke, Michael Xie, W Matthew Davis, David B Lobell, and Stefano Ermon. Combining satellite imagery and machine learning to predict poverty. *Science*, 353(6301):790–794, 2016.

Morten Jerven. *Poor numbers: how we are misled by African development statistics and what to do about it*. Cornell University Press, 2013.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *Proceedings of the 4th International Conference on Learning Representations (ICLR)*, 2015.

National Institute of Statistics of Rwanda (NISR). Poverty trend analysis report Rwanda, 2016.

Anthony Perez, Christopher Yeh, George Azzari, Marshall Burke, David Lobell, and Stefano Ermon. Poverty prediction with public landsat 7 satellite imagery and machine learning. *arXiv preprint arXiv:1711.03654*, 2017.

Shea Oscar Rutstein and Kiersten Johnson. *The DHS wealth index*. ORC Macro, 2004.

Nils B Weidmann and Sebastian Schutte. Using night light emissions for the prediction of local wealth. *Journal of Peace Research*, 54(2):125–140, 2017.