LEARNING TO ESTIMATE POVERTY FROM SPACE : USING SATELLITE IMAGERY AND DEEP LEARNING TO CREATE TEMPORAL POVERTY MAPS FOR SOUTH AFRICA

Keon M.L. Roohani

School of Electrical & Information Engineering, University of the Witwatersrand, Private Bag 3, 2050, Johannesburg, South Africa

Abstract: Policy makers, investors, and charity organisations require accurate, repeatable, and low-cost techniques to identify areas of extreme poverty throughout the world if the UN's goal to end poverty worldwide by 2030 is to be achieved. This investigation employs and reproduces a known method which uses satellite imagery, at the village or neighbourhood level, and a convolutional neural network to predict asset wealth estimates derived using in-person surveys throughout Africa. The method was extended to the South African context showing that a ResNet-18 architecture trained on multi-spectral and nightlight image bands can explain up to 60% of the spatial variation in asset wealth. The model was then used to generate temporal poverty maps of South Africa from 2016 to 2021 which are shown to correlate at the macro-level with changes in economic indicators, such as GDP, over time. Potential future work is discussed which involves using additional data sources to improve results and extending the methodology to other indicators of human well-being.

Key words: Computer Vision, Deep Learning, Machine Learning, Poverty, South Africa, UN Sustainable Development Goals.

1. INTRODUCTION

The United Nations (UN) has put forward 17 sustainable development goals to be achieved by the year 2030 [1]. The first of these goals is to end poverty everywhere [1]. This goal cannot be accomplished unless and until policy makers, investors, and charity organisations are provided with an accurate, repeatable, and low-cost means to identify areas of extreme poverty throughout the world. Recent advancements in remote sensory imaging have enabled the use of satellite imagery as a proxy for standard census and survey information [2, 3]. The aim of this investigation was twofold. Firstly, to utilise and reproduce the existing methods described in the work of Yeh et. al. [3] to predict poverty for 23 countries in sub-Saharan Africa. These methods involve using multi-spectral satellite imagery and deep learning to predict poverty at the village or neighbourhood level throughout Africa. The second aim was to extend these methods to the South African context in order to produce temporal poverty maps which display the changing economic landscape of South Africa over time and space. Section 2 provides a background of existing methods used to identify well-being. Section 3 describes the use of known methods to predict well-being throughout Africa. Section 4 details improvements which allow well-being to be mapped solely within South Africa. Finally, section 5 provides a discussion and analysis of the results received during the investigation and highlights possible improvements.

2. BACKGROUND

2.1 Methods for Measuring Human Well-Being

The ability to accurately measure human well-being forms an imperative part of the United Nations' goal to end poverty everywhere [1]. Without sufficient socio-geographical data, and corresponding economic indicators, the necessary steps required to alleviate financial strain on the poor cannot be performed. This is of particular concern in sub-Saharan Africa where reports as recent as 2021 indicate that roughly 37% of people live on less than \$1.90 a day, which is considered abject poverty by the UN [1]. The two main methods used to provide policy makers with relevant information are in-person asset wealth or consumption surveys, and modern predictive satellite imaging techniques [2, 3]. Both methods have their advantages and disadvantages with regards to feasibility, frequency, and accuracy.

2.1.1 Classical Survey Methods: Traditionally, required socio-economic information has been acquired using in-person surveys. This is an expensive and time consuming process which is, thus, infrequently performed. These surveys collect information regarding individual or household health, poverty, education, and other living standards markers [4, 5]. These markers are then used to construct comparable 'asset wealth' indices using a standardized principal component analysis (PCA) [4]. Asset wealth indices have shown to perform better than income and expenditure as predictors of wealth in low and middle income countries [5]. Unfortunately, despite global effort through initiatives, such as the Demographic and Health Survey (DHS) Program, the World Bank Community Surveys, and national population censuses, African countries undertake surveys, on average, less frequently than every four years [3]. This, coupled with a lack of temporal consistency due to different households being questioned in subsequent surveys, has given rise to the need for modern techniques which utilise existing survey data to predict wealth in unsurveyed locations.

2.1.2 Modern Satellite Imaging Methods: The high cost and low frequency of in-person survey data has made data collected from other sources, such as cellphone networks, social media, and satellite imagery, an appealing proxy with which to identify poverty [2]. Initial attempts made use of "nighlight" luminosity data to predict well-being with some success [6, 7]. However, the greatest strides have come from the adoption of daytime imagery, especially those which make use of multi-spectral bands, to predict economic well-being [2, 3]. The data is taken at varying resolutions from satellites, whose images lie in the public domain, and are trained using convolutional neural networks (CNNs) to provide predictions of economic well-being across both time and space. However, these methods are only viable assuming that the error due to using satellite imagery as a predictor of well-being is uncorrelated with the standard error found in classical surveys [7]. Under this assumption, the use of multi-spectral satellite imagery to train deep learning models has been shown to describe up to 70% of the variation in asset wealth at the village or neighbourhood level [3].

2.2 Wealth and Poverty Throughout Africa

Africa suffers the highest extreme poverty rates in the world [1]. It consists primarily of developing low and middle income countries. This makes accurate and frequent measurement of well-being a primary concern for policy makers and investors. Despite the plethora of poor countries found on the continent, it is also home to countries such as South Africa which are notably wealthier than their neighbours. South Africa finds itself in a unique socio-economic position, with striking extremes of wealth and poverty apparent throughout its landscape. These truths make it crucial to ensure that existing poverty identification methods used throughout Africa, such as those pioneered by Yeh et. al. [3], are transferable to the South African context.

3. PREDICTING WELL-BEING THROUGHOUT AFRICA

3.1 Methodology and Experiments

The methodology utilised to predict wealth throughout Africa is comprehensively described in the work of Yeh et. al. [3], however the salient points are described here for completeness. Furthermore, the addition of relevant data for South Africa is described. Two crucial datasets are required before deep learning models can be trained to predict well-being.

The first is comprised of the asset wealth indices determined using DHS program survey data spanning from 2009 to 2016. These are computed using a standardized PCA which provides a single asset wealth value based on the quality of a set of assets such as toilets, flooring, utensils, and other valuables [3]. A summary of the pre-computed asset wealth data for the 23 sub-Saharan African countries is shown in figure 1. This scale ranges from -1.3 to 2.6 chosen to map the mean asset wealth to 0. Notably, the data for the 2016 DHS survey of South Africa has been added, however, a different but comparable asset wealth index was used in this case. This is known as the International Wealth Index (IWI) which was created using data from over 2 million households in low and middle income countries [5]. The IWI identifies asset wealth on a scale of 0 - 100, but was normalised to fit the scale of the pre-computed values used in the work of Yeh et. al. [3]. The initial dataset excluding South Africa is comprised of 19669 clusters, where clusters are defined as villages in rural areas or neighbourhoods in urban areas [3]. The additional data for South Africa consists of 746 clusters from the 2016 DHS survey. The geographic positioning of each cluster is displaced by roughly 2km for urban areas and up to 10km for rural areas, meaning that there is sufficient noise in the data to maintain privacy.

The second required dataset is comprised of the satellite images for each DHS cluster. The images were obtained from the Landsat 5, Landsat 7, and Landsat 8 satellites. These satellites possess both daytime and nightlight images. The nightlight images consists of a DMSP [8] band for the 2009 - 2011 data, and a Viirs [9] band for the 2012 - 2016 data. The daytime imagery contains seven multi-spectral bands, these include: The standard Red, Green, and Blue bands which form an RGB image, a Near Infrared (NIR) band, two Shortwave Infrared bands (SWIR1 & SWIR2), and a Thermal band (TEMP1). Each band highlights unique aspects of the satellite imagery such as colour, vegetation, temperature, and rock formations. All images are 255 x 255 pixels and have a resolution of 30 meters per pixel. The final dataset was formed by taking the pixel average of many cluster images over a given time period in order to minimize the effect of cloud cover.

These data points were used as the input to a CNN with a ResNet-18 architecture implemented using Python and Tensorflow. The satellite images were center cropped to 224 x 224 pixels and paired with a corresponding asset wealth index value in order to match the input shape of the ResNet-18 model.

The experiments were performed on a combination of satellite image bands including multi-spectral (MS) only, nightlight (NL) only, and multi-spectral with nightlights (MS + NL). Each experiment was performed five times with each model created on a different fold, or subset of the data with its own training, validation, and testing split. The five models were then merged, fine tuned, and cross validated using ridge regression on unseen data in order to reduce overfitting of the final model. There are two main types of experiments: In-country (IC) and out-of-country (OOC) experiments. For OOC experiments an entire country forms part of either the training, validation, or testing split; this validates the model's performance on unseen countries. Whereas for IC experiments all splits contain some portions of each country providing a more varied set of splits. In order to maintain

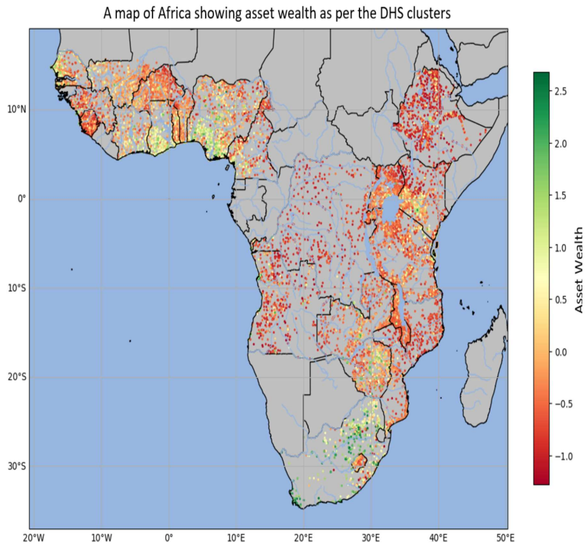

Figure 1: A map of Africa showing the DHS cluster locations and corresponding asset wealth index values for the 23 sub-Saharan African countries from 2009 to 2016 used to train the deep learning models. The 2016 DHS data available for South Africa and corresponding normalised IWI values are also included.

uniformity, each model was trained for 50 epochs on the 'WISE01' computer which contains an NVIDIA RTX 2080 Ti graphics card and 128 gigabytes of RAM. The experiments were performed on the 19669 DHS clusters excluding South Africa and then repeated on the 20415 clusters with South Africa included.

3.2 Results

The experiments performed on the 19669 clusters across 23 sub-Saharan African countries were compared against the results obtained by Yeh et. al. [3] as shown in Table 1. The r^2 value is the average variance of the model correlation and the MSE is the mean squared error of the experiment predictions. The higher the r^2 value the better the performance of the model, however an r^2 very close to 1 may be a sign of over-fitting of the model. The lower the MSE the more accurate each individual prediction; the MSE is bounded by the asset wealth scale used in figure 1. The multi-spectral with nightlight in-country experiment describes the variation in wealth most accurately with the lowest mean squared error (MSE) of 0.202. The multi-spectral only experiment, using out-of-country splits, performed the worst with an MSE of 0.293. This is congruent with expected results as seen in Table 1 and provides confidence in the implementation of the framework. Each model trained in our investigation performed similarly to its original counterpart.

The experiments were repeated with South Africa added to the datasets as shown in figure 1. This resulted in 20415 available clusters used in each experiment. The results of these experiments were similar to those shown in table 1, meaning that the addition of South Africa did not notably negatively affect the average model performance. However, the results of South Africa individually were substantially below the norm. The best performing model for the 746 South African clusters had an $r^2 = 0.304$ and a comparably high MSE = 0.699. This means that directly extending the methodology used in the original paper by Yeh et. al. [3] to the South African context did not produce usable results. The reason for this can be found in the anomalous nature of South Africa as a nation, as it is comparably wealthier than its neighbours. This is evidently witnessed in figure 1 with South Africa showing many 'greener' (wealthier) clusters. The models trained on data throughout Africa predict wealth in South Africa significantly lower than its true value. This is especially true of the urban areas, while the poorer rural areas are slightly more accurately modelled. To address these poor results alternative methods to map well-being solely throughout South Africa have been proposed below.

Table 1: The results of the ResNet-18 experiments performed on the 23 sub-Saharan African countries which utilise satellite imagery to predict asset wealth. Columns marked 'OR' represent the original results obtained by Yeh et. al. [3].

Experiment	r^2	MSE	r^2 (OR)	MSE (OR)
MS+NL OOC	0.656	0.228	0.668	0.217
NL Only OOC	0.641	0.246	0.661	0.221
MS Only OOC	0.588	0.293	0.615	0.252
MS+NL IC	0.696	0.202	0.703	0.195
NL Only IC	0.655	0.229	0.681	0.208
MS Only IC	0.671	0.214	0.674	0.214

4. PREDICTING WELL-BEING SOLELY WITHIN SOUTH AFRICA

4.1 Methodology and Experiments

In an attempt to improve the accuracy of the models used to predict well-being throughout South Africa some alterations were made to the methodology described in section 3.1. Rather than using all 20415 clusters, only the 746 clusters from the 2016 DHS program survey were used as the dataset for the experiments. This dataset represents 11083 households throughout South Africa, which is on average 14.85 households per cluster. The geographical and wealth distribution of these clusters is given in figure 2. Unfortunately, no other DHS program surveys have been conducted within the last fifteen years. While there are other sources of survey data available, they are not suited to the methodology explored in this investigation. The Community Survey 2016 and Census South Africa 2011 were explored, however they were found to be unsuited to the chosen methodology due to only providing data at the municipal level and not always providing sufficient asset wealth information.

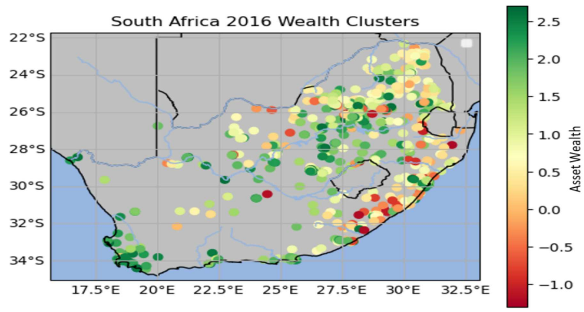

Figure 2: A map of South Africa showing the cluster locations and corresponding IWI values for the 2016 DHS program survey.

The dataset was split into five folds as before, however, in these experiments the training, validation, and testing splits for each fold consisted of different subsections of the South Africa data assigned to ensure no overlap of image points. Any overlap in the area covered by the satellite images would distort the results of the final model.

The experiments performed made use of a similar CNN with a ResNet-18 architecture. They were performed on the MS, NL, and MS + NL bands as discussed previously. However, in this case there is no distinction between an in-country and out-of-country experiment. Additionally, each model was trained to 500 epochs rather than 50 to accommodate for the reduced dataset size.

4.2 Results

The results obtained from the experiments using only South African data were distinctly better than the approach which included all of Africa. Even the worst performing model, which was the MS only experiment, had an $r^2 = 0.344$ and a MSE = 0.593 which outperformed the original experiments. The NL only experiment obtained a slightly superior $r^2 = 0.449$ and an MSE = 0.498. Finally, the best performing model by far was the MS + NL model which obtained an $r^2 = 0.600$ with an MSE = 0.361. The MS + NL model performance is seen in figure 3.

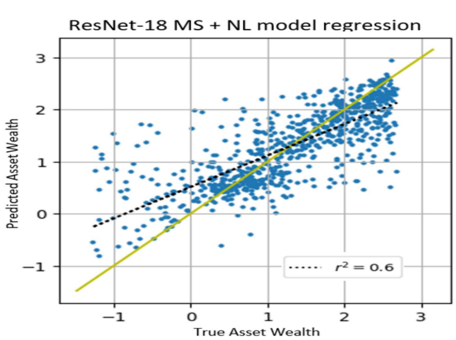

Figure 3: The multi-spectral (MS) with nightlight (NL) model trained on South Africa only showing the correlation between predicted and true IWI values.

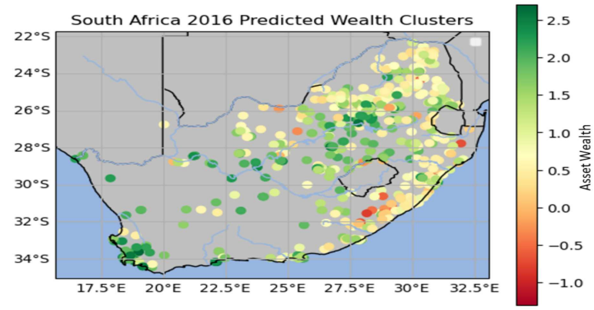

Figure 4: A map of South Africa showing the cluster locations and corresponding predicted IWI values for the year 2016 using the MS + NL ResNet-18 model.

While the MS + NL model only describes 60% of the variation in wealth across South Africa, this is a stark improvement over the original 30.4% shown using the previous approach. This model differentiates adequately between the extremes of wealth and poverty present within South Africa. A display of the predictions can be seen in figure 4. When contrasting the true IWI values, shown in figure 2, and the predicted values for the same year, shown in figure 4, certain observations regarding the strengths and weaknesses of the model can be made. Firstly, the model distinguishes between larger regions of wealth and poverty correctly, with the Western Cape and Gauteng correctly identified as generally wealthier than regions such as Limpopo, the Eastern Cape, and KwaZulu-Natal. Secondly, the model tends to consider areas of extreme poverty as slightly wealthier than their true value. This is a weakness most likely due to the IWI wealth distribution of South Africa being biased towards generally wealthier values. Thirdly, the model struggles to differentiate between extremes of wealth and poverty within a small region. For instance, Gauteng consists of both highly wealthy urban areas as well as poorer communities which include informal settlements. However, the model does not highlight these disparities as clearly as is seen in the survey calculated IWI values of figure 2.

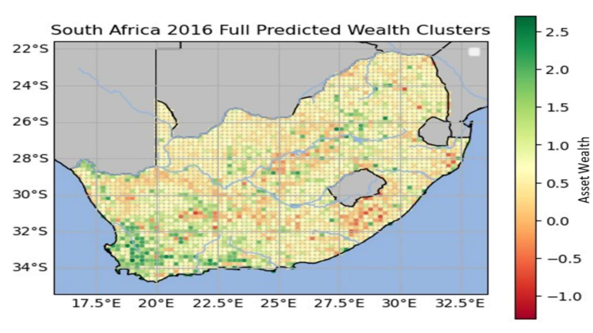

Figure 5: A map of South Africa showing 2840 uniformly distributed satellite images and corresponding predicted IWI values for the year 2016 using the MS + NL ResNet-18 model.

To provide a clearer understanding of the model's

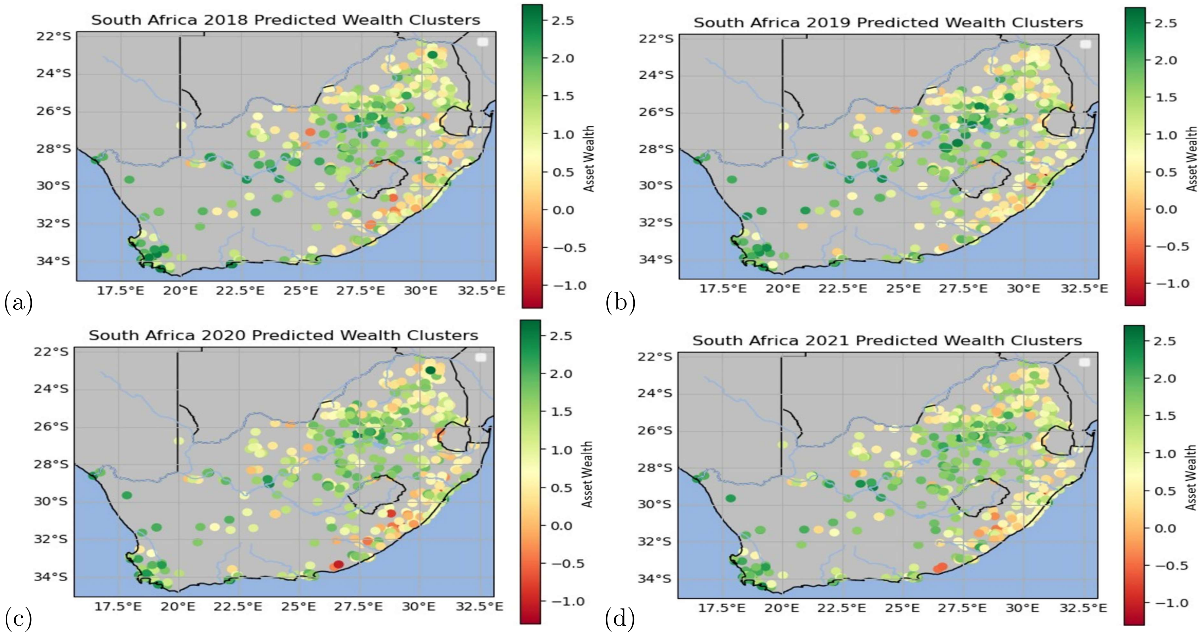

Figure 6: IWI predictions for South Africa using the MS + NL ResNet-18 model for satellite imagery downloaded for the years (a) 2018, (b) 2019, (c) 2020, and (d) 2021

performance a set of 2840 satellite images, uniformly distributed throughout South Africa, were sourced and displayed as seen in figure 5. Despite the images being of locations different to the DHS survey, the resulting predictions show a very similar distribution of wealth when compared to figure 4. This provides further confidence in the ability of the model to differentiate wealth spatially.

In order to verify the MS + NL ResNet-18 model's ability to explain temporal changes in wealth, figure 6 displays the predictions for the 746 DHS clusters for the years 2018, 2019, 2020, and 2021. Clearly there is some variation in each individual prediction, however, to validate the temporal correctness against known DHS values from years other than 2016 is impossible due to a lack of survey data. In an attempt to remedy this quandary, South Africa's yearly GPD was used as an indicator of household wealth. This is a fair assumption as it has been shown that there is correlation between household income and GDP, and by extension household asset wealth and GDP [10].

A rise in the overall GDP per capita should translate to an increase in the mean standard of living across South Africa. The mean IWI predictions of the model were mapped against the known GDP values for each year as seen in figure 7. Both the predicted IWI and the GDP values follow the same trend. This is an increase from 2016 to 2018, followed by a decrease from 2018 to 2020, followed by a large increase in 2021. This major dip in both GDP and mean IWI around 2020 can be explained by the COVID-19 pandemic and the economic strains it placed on South Africa. This shows that the model's predictions vary at the macro-level

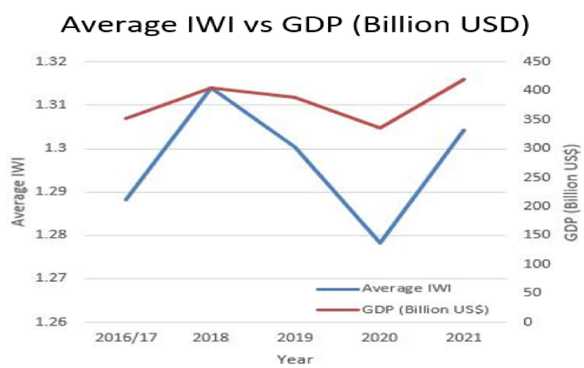

Figure 7: The known Gross Domestic Product (GDP) of South Africa from 2016 to 2021 mapped against the mean IWI predictions of the ResNet-18 MS + NL model for the same years.

in line with other known economic indicators, showing the model has at least some sensitivity to changes of a temporal nature.

5. DISCUSSION AND ANALYSIS

The method of using satellite imagery to predict economic well-being, pioneered by Yeh et. al. [3], has been reproduced showing scalable and accurate performance on unseen countries throughout Africa. Unfortunately, this high performance does not translate to South Africa when included in the experiments, likely due to South Africa's generally wealthier economic status. Thus, the method has been adapted to identify poverty solely in South Africa with an increased prediction performance of the spatial variation in household asset wealth. This new method suffers from a lack of data points, both spatial and temporal. This could be improved by grouping countries with similar country-wide wealth levels and socio-geographic distributions of wealth and poverty. For example, South Africa could be investigated alongside similar African, South American, or Asian middle income countries in order to increase the data available for training. The methodology could also be improved to increase the final model's ability to differentiate between extremes of wealth and poverty in a small geographical region. This could be done by using higher resolution satellite imagery of less than 30 meters per pixel. Furthermore, the temporal validity of the models could be improved by repeating these experiments once newer DHS data is released. Despite this lack of temporal data, the results show a sensitivity between predicted asset wealth and other macro-level economic indicators. This is a sign that the methodology is successful in its aim to measure temporal changes in economic well-being. The methodology could be extended and used to predict other indicators of wealth such as income or consumption, or even other welfare indicators such as health and education, assuming there are sufficient data points. Clearly, these satellite image based methods do not serve as a replacement to traditional surveys, but rather supplement them and fill in the gaps where data is unavailable.

6. CONCLUSION

Satellite imagery has been used to predict economic well-being throughout sub-Saharan Africa describing up to 70% of the variation in asset wealth. These methods have been extended to South Africa and have provided models that explain up to 60% of the variation in asset wealth when trained using multi-spectral and nightlight satellite image bands. Predicting temporal changes in asset wealth throughout South Africa has been a challenge due to a lack of high quality data, however, the final models show sensitivity to changes in macro-level economic indicators such as gross domestic product. Potential future work and improvements to the chosen methodology have been described which mainly involve finding ways to increase available high quality data sources. This investigation has generated temporal poverty maps for South Africa which should assist policy makers in combating poverty and achieving the UN's goal to eradicate poverty by 2030.

ACKNOWLEDGEMENTS

Acknowledgements are given to the Demographic and Health Survey (DHS)[1] Program and Global Data Lab[2] for access to their survey and wealth index information. Further thanks are given to Google Earth Engine[3] for providing free access to its satellite image database. Acknowledgements are also given to Yeh et. al. [3] for access to their code for this investigation[4]. Finally, many thanks to Professor Turgay Celik who supervised this investigation and provided insightful assistance. Additionally, access to the 'WISE01' computer was much appreciated, as it expedited the training and processing of the deep learning models.

[1]https://dhsprogram.com/
[2]https://globaldatalab.org/
[3]https://earthengine.google.com/
[4]https://github.com/sustainlab-group/africa_poverty

References

[1] United Nations, "Progress towards the Sustainable Development Goals," *Report of the Secretary-General*, 2022.

[2] N. Jean, M. Burke, M. Xie, W. M. Davis, D. B. Lobell, and S. Ermon, "Combining satellite imagery and machine learning to predict poverty," *Science*, vol. 353, no. 6301, pp. 790–794, 2016. DOI: 10.1126/science.aaf7894.

[3] C. Yeh *et al.*, "Using publicly available satellite imagery and deep learning to understand economic well-being in Africa," *Nature Communications*, vol. 11, no. 1, May 2020, ISSN: 2041-1723. DOI: 10.1038/s41467-020-16185-w.

[4] D. E. Sahn and D. Stifel, "Exploring alternative measures of welfare in the absence of expenditure data," *Review of Income and Wealth*, vol. 49, no. 4, pp. 463–489, 2003. DOI: https://doi.org/10.1111/j.0034-6586.2003.00100.x.

[5] J. Smits and R. Steendijk, "The International Wealth Index (IWI)," *Social Indicators Research*, vol. 122, no. 1, pp. 65–85, 2015, ISSN: 1573-0921. DOI: 10.1007/s11205-014-0683-x.

[6] T. Ghosh, S. J. Anderson, C. D. Elvidge, and P. C. Sutton, "Using nighttime satellite imagery as a proxy measure of human well-being," *Sustainability*, vol. 5, no. 12, pp. 4988–5019, 2013, ISSN: 2071-1050. DOI: 10.3390/su5124988.

[7] J. V. Henderson, A. Storeygard, and D. N. Weil, "Measuring economic growth from outer space," *American Economic Review*, vol. 102, no. 2, pp. 994–1028, 2012. DOI: 10.1257/aer.102.2.994.

[8] F.-C. Hsu, K. E. Baugh, T. Ghosh, M. Zhizhin, and C. D. Elvidge, "Dmsp-ols radiance calibrated nighttime lights time series with intercalibration," *Remote Sensing*, vol. 7, no. 2, pp. 1855–1876, 2015, ISSN: 2072-4292. DOI: 10.3390/rs70201855.

[9] C. D. Elvidge, K. Baugh, M. Zhizhin, F. C. Hsu, and T. Ghosh, "Viirs night-time lights," *International Journal of Remote Sensing*, vol. 38, no. 21, pp. 5860–5879, 2017. DOI: 10.1080/01431161.2017.1342050.

[10] O. Causa, S. Araujo, A. Cavaciuti, N. Ruiz, and Z. Smidova, "Economic growth from the household perspective," no. 1111, 2014. DOI: https://doi.org/https://doi.org/10.1787/5jz5m89dh0nt-en.