Constructing a Visual Dataset to Study the Effects of Spatial Apartheid in South Africa

Raesetje Sefala
DAIR Institute

Timnit Gebru
DAIR Institute

Luzango Mfupe
CSIR

Nyalleng Moorosi
Google Research

RIchard Klein
Wits University

Intro & Motivation

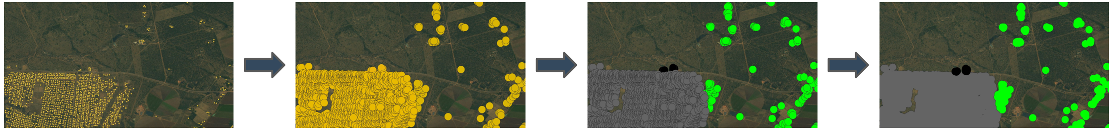
Removing many of the legacies of apartheid is an important problem in South Africa.

Figures 1 and 2 show some aerial images taken by photographer Johnny Miller, depicting completely segregated neighbourhoods of townships next to gated wealthy neighbourhoods that have largely remained unaffected by the ending of apartheid [1].

Studying changes in the demographic makeup of different neighborhoods could help implement policies to desegregate them.

This paper introduces the first publicly available dataset to study the evolution of spatial apartheid.

We describe our iterative process to create this dataset over two years, which includes pixel wise labels for 4 classes of neighborhoods: wealthy areas, non wealthy areas, nonresidential neighborhoods and background (undeveloped land).



Dataset Creation

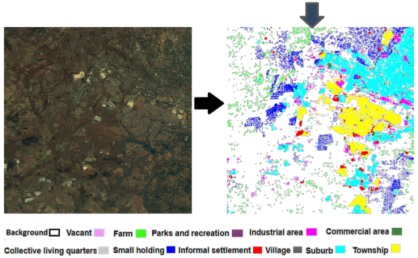
To assemble our dataset we use 3 publicly available datasets depicting South Africa in 2011:

*High resolution satellite images
*Geo-referenced buildings dataset: points denoting where all the buildings are in South Africa
*Enumeration Areas: polygons denoting South Africa's land use labels as designated by the government

**township class was hand-labelled

1. Using the building dataset, find the building points.
2. Expand the points to cover the houses by buffing.
3. Intersect the expanded building data with the land use classes.
4. Smoothing overlapping building polygons by neighborhood type.
5. 12 Classes of neighborhood types across the entire country.

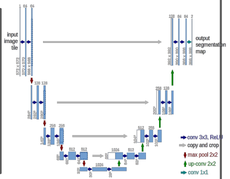
We collapsed classes into
background (all land without buildings)
wealthy neighborhoods (suburbs, smallholdings, and farms)
non wealthy neighborhoods (townships, informal settlements, villages and collective living quarters)
non residential building clusters (commercial areas, industrial areas, buildings on vacant land, parks, and recreational areas)

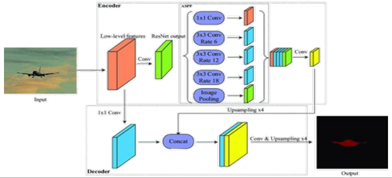


Models

*Sampled 1,869,840 images from the 10,284,120 images covering the country.

*Trained a UNet [2] (right) and a Deeplab [3] (below) model on the datasets.





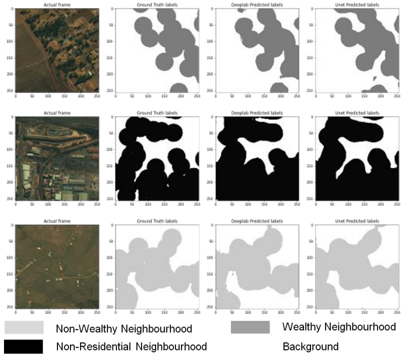
Results

*Our best model (UNet) on 12 classes of neighborhood types: MIoU = 45.85%.

*Comparison: SOTA on 10 Classes of neighborhood types had MIoU = 54.96% [4].

*Our best model (UNet) on 4 classes of neighborhood types: MIoU = 66.22%.

*The UNet performed the best on both 12 and 4 classes on all experiments.



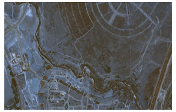
Conclusion & Figure work

We have introduced the first visual dataset of South Africa which can be used to analyze the effects of spatial apartheid, and described our iterative data annotation process that allowed us to assemble this dataset.

*Future work we expect to use this dataset for includes exploring how neighborhoods have changed in different parts of the country.

*The images on the right show how we use the model trained on 2011 images ((a) & (b)) to estimate the changes in 2017 ((c) &(d)).

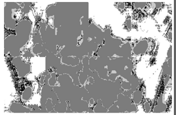
With our dataset, we hope to enable those interested in studying and reversing the effects of spatial apartheid, to use this dataset.



(a) 2011 Satellite Image

(b) 2011 ground truth

(c) 2017 Satellite Image



(d) 2017 Model Predictions

References

[1] Johnny Miller. Apartheid's urban legacy, in striking aerial photographs. https://bit.ly/ 3bB3fbo, 2016 accessed January 7, 2020.
[2] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, pages 234–241. Springer, 2015.
[3]Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam. Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European conference on computer vision (ECCV), pages 801–818, 2018.
[4] Yan Liu, Qirui Ren, Jiahui Geng, Meng Ding, and Jiangyun Li. Efficient patch-wise semantic segmentation for large-scale remote sensing images. Sensors, 18(10):3232, 2018.

DEEP LEARNING INDABA

Acknowledgements

Thank you to Alex Hanna, Ben Packer, Deb Raji, David Everatt, Danielle Wood, Memory Mhembere, Samy Katumba, Emily Denton, Neil Lawrence, Xiao Wei for insightful discussions and suggestions. Funding for this project was provided through five sources: The South African Department of Science and Technology and the Council of Scientific & Industrial Research (as a masters scholarship award to Raesetje Sefala), Google (as a research award and compute credits to Raesetje Sefala), the Deep learning Indaba and Nvidia (Nvidia Titan V prize for best poster presentation at the 2018 Deep Learning Indaba summer school), and the Distributed AI Research Institute (DAIR).