

UNIVERSITY OF MÜNSTER  
DEPARTMENT OF INFORMATION SYSTEMS

---

Thesis Exposé  
Deep Learning Semantic Segmentation of Tree  
Stock in South Africa Using Satellite Images

---

BACHELOR THESIS

submitted by

Leo Richard Hermann Giesen

CHAIR OF DATA SCIENCE:  
MACHINE LEARNING AND DATA ENGINEERING

<b>Principal Supervisor</b>	PROF. DR. FABIAN GIESEKE
<b>Supervisor</b>	SUPERVISOR, SUBST.-PROF. DR. FRIEDRICH CHASIN Chair for Data Science: Machine Learning and Data Engineering
<b>Matriculation Number</b>	462502
<b>Field of Study</b>	Information Systems
<b>Contact Details</b>	leo.giesen@uni-muenster.de
<b>Submission Date</b>	12.01.2022

---

## Motivation

The bachelor thesis builds upon Brandt et al. [Bra+20], which outlined a relatively high density of isolated trees and shrubs (hereafter collectively referred to as trees) in the Sahel, which challenges the desertification of central Africa. The detection of trees in this research project enables an estimate of biomass and corresponding CO<sub>2</sub> emissions. Many trees grow isolated without canopy closure and are not part of a larger forest. However, these non-forest trees crucially contribute to the drylands' biodiversity and ecosystem because they store carbon, offer food resources, and shelter humans and animals [Bra+20; Str+12; Bay+14].

Unfortunately, there has not been sufficient public research concerning isolated tree coverage [SKS15]. Therefore, this thesis aims to check if this insight concerning desertification is transferable to South Africa with an independent implementation using the state-of-the-art technology U-Net. The provided satellite image data spans over several years. Hence, the tree stock and vegetation development is observed and may be analyzed by local authorities or environmental initiatives. The resulting map of vegetation could be presented on a website to create awareness for South Africa's deforestation and may be extended to its land loss.

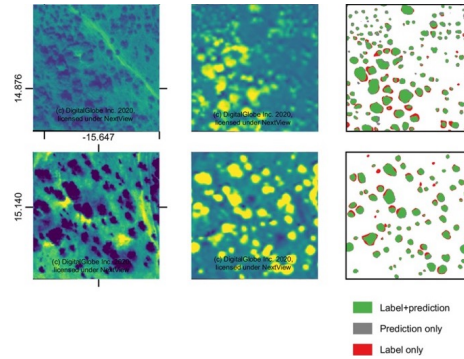


Figure 1 Tree Mapping with U-Net [Bra+20]

## Problem Statement & Potential Outcome

First and foremost, the first objective is to investigate to what extent Deep Learning architectures such as U-Net are suitable for detecting trees with the present data. According to Brandt et al. [Bra+20], it is possible to create "excellent models for detecting isolated trees over large areas" with a Deep Learning approach combined with very high-spatial-resolution satellite image data, which is the case for the present data because it offers an extremely high resolution of 25 cm per pixel. However, it is also pointed out that the "transferability of the model across regions can be low." This

---

may be due to a significant difference in tree landscape characteristics, which affect the algorithm's ability to differentiate trees from the surrounding landscape. Thus, a comparison of Sahel's and South Africa's landscape aids in estimating the model's transferability. The training data from Brandt et al. [Bra+20] explored drylands, deserts, and semi-arid grassland [Bra+20; SA04]. Upon first inspection and discussion with Thorben Hellweg, the satellite images cover Limpopo and Mpumalanga provinces, i.e. the Kruger National Park. The landscape is partly mountainous, bush plains and tropical forests [Nat15].

Furthermore, the second and central objective is the development of a Deep Learning model, which detects vegetation in the South African drylands, specifically trees and shrubs with a minimal crown size of three meters. The Convolutional Neural Network architecture U-Net is to be utilized, which is built for "fast and precise segmentation of images" [RFB15]. Based on the learnings of Brandt et al. [Bra+20], specific suitable characteristics may be adopted for this project. For instance, the Tversky loss function or the usage of batch normalization layers after each convolution block. The mapping of trees results in various metrics affecting sustainability, such as tree density, crown size, and tree cover. A map of predicted trees in QGIS will provide an overview of tree coverage, which will be compared and evaluated with other sources such as Hansen et al. [Han+13]. Brandt et al. [Bra+20] required a vast number of training samples to perform well. Due to a tight schedule, this project will not produce as many training instances. This problem could be solved by Transfer Learning or generating new training instances by correcting the model's prediction of unlabeled areas. Finally, hyperparameter grid search could be performed to enhance the model's performance. The applied software comprises data labeling in QGIS (3.22.1) and Deep Learning model is implemented in Python (3.9.5) using TensorFlow (2.6.1) and possibly Keras (2.4.0).

## **Work Plan & Research Approach**

First, the data is accessed on the university server using the remote desktop gateway and is explored and labeled in QGIS. Second, data preparation and the Deep Learning implementation are performed in a Jupyter Notebook. Third, the U-Net model is chosen. Fourth, the model is trained and evaluated. Fifth, the application of the model is performed by making a prediction. Seventh, research about U-Net and the project's relevance in a sustainability context with an Information System application is conducted throughout the project. Eighth, the structure of the thesis is developed once all significant insights from the implementation have been made. Ninth, the thesis is formulated and revised.

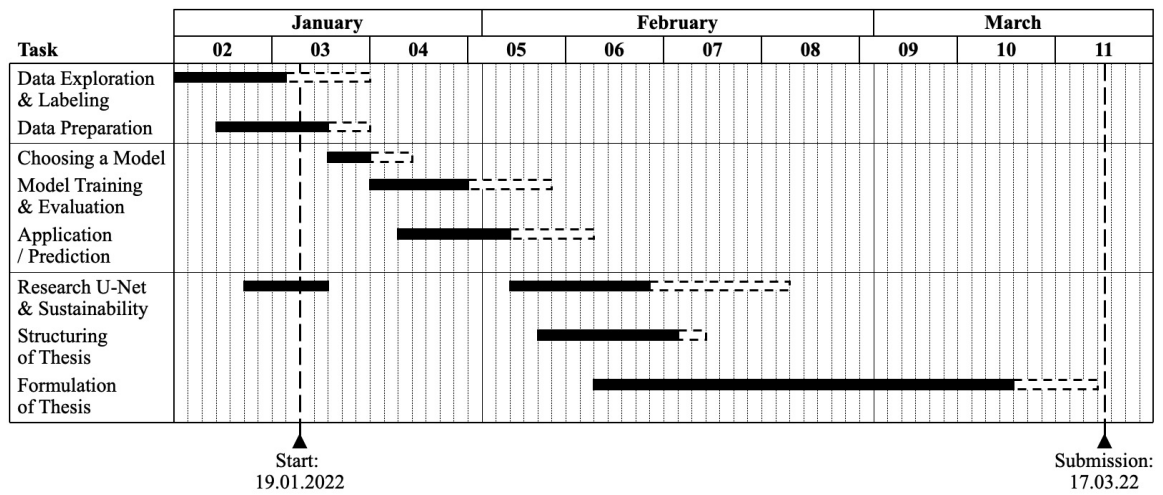


Figure 2 Project Schedule

## Bibliography

- [Bay+14] J. Bayala et al. “Parklands for Buffering Climate Risk and Sustaining Agricultural Production in the Sahel of West Africa”. In: *Current Opinion in Environmental Sustainability* 6.1 (Feb. 2014), pp. 28–34.
- [Bra+20] Martin Brandt et al. “An Unexpectedly Large Count of Trees in the West African Sahara and Sahel”. In: *Nature* 2020 587:7832 587.7832 (Oct. 2020), pp. 78–82.
- [Han+13] M. C. Hansen et al. “High-Resolution Global Maps of 21st-Century Forest Cover Change”. In: *Science* 342.6160 (Nov. 2013), pp. 850–853. URL: <https://earthenginepartners.appspot.com/science-2013-global-forest>.
- [Nat15] National Geographic Society. *South Africa*. 2015. URL: <https://kids.nationalgeographic.com/geography/countries/article/south-africa> (visited on 12/29/2021).
- [RFB15] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. “U-Net: Convolutional Networks for Biomedical Image Segmentation”. In: *Lecture Notes in Computer Science at the University of Freiburg* 9351 (2015), pp. 234–241.
- [SA04] Robert Simmon and Jesse Allen. *Vegetation and Rainfall in the Sahel*. 2004. URL: <https://earthobservatory.nasa.gov/images/7277/vegetation-and-rainfall-in-the-sahel> (visited on 12/29/2021).
- [SKS15] Sebastian Schnell, Christoph Kleinn, and Göran Ståhl. “Monitoring Trees Outside Forests: A Review”. In: *Environmental Monitoring and Assessment* 2015 187:9 187.9 (July 2015), pp. 1–17.
- [Str+12] L. C. Stringer et al. “Challenges and Opportunities in Linking Carbon Sequestration, Livelihoods and Ecosystem Service Provision in Drylands”. In: *Environmental Science Policy* 19-20 (May 2012), pp. 121–135.