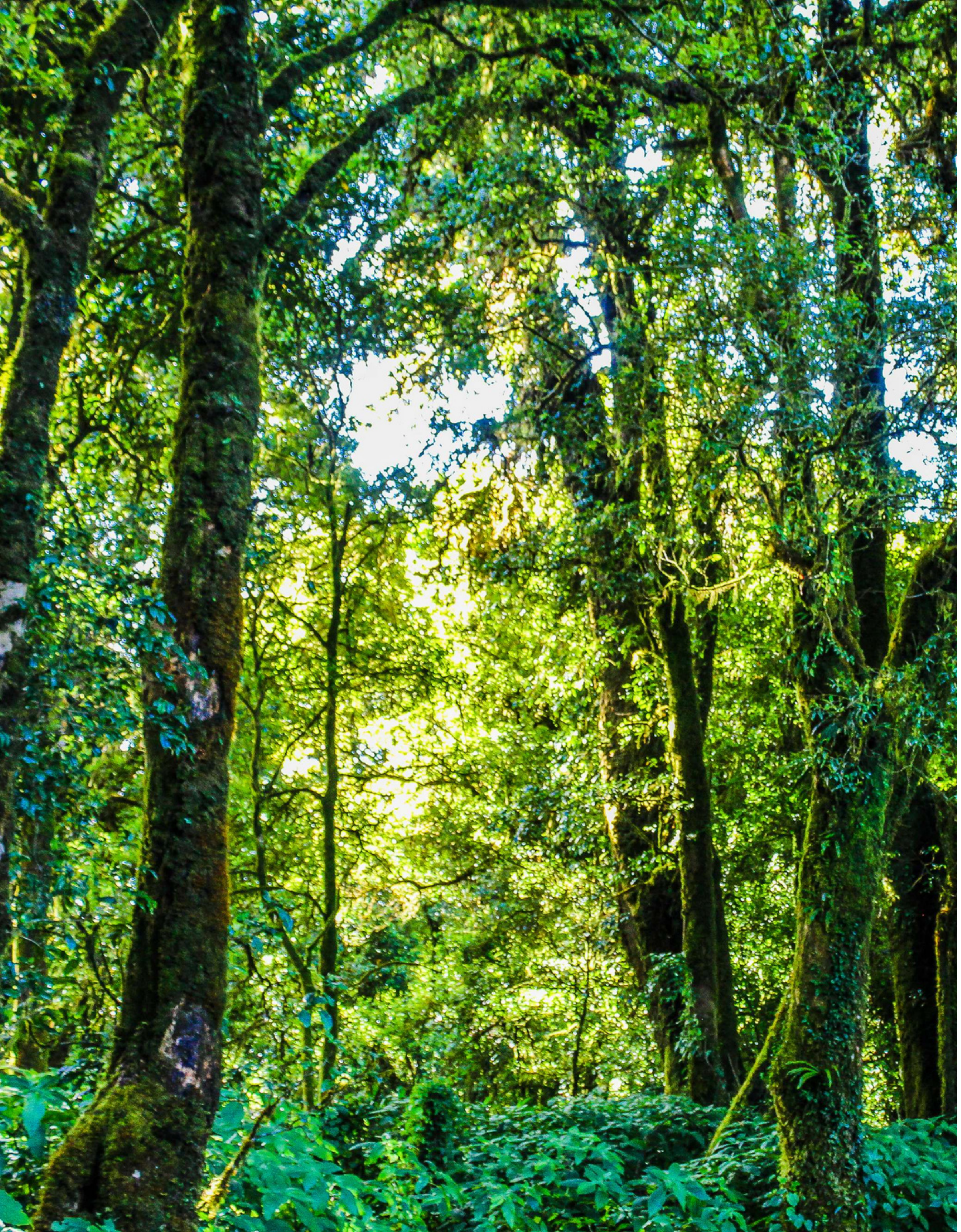


Technische Hochschule Würzburg-Schweinfurt

Weakly Supervised Semantic Segmentation of Multispectral Satellite Imagery for Land Cover Mapping

Esther Ademola
25.03.2024

Agenda

- Introduction
- Objective
- Research Questions
- Related Works

Introduction

Climate change poses a pressing global challenge, impacting ecosystems and weather patterns. Deforestation, driven by human activities, exacerbates these issues by altering landscapes and intensifying climate change effects.

Objective

Technische Hochschule Würzburg-Schweinfurt

The primary aim of this thesis is to leverage weakly supervised semantic segmentation techniques on high-resolution multispectral data to achieve accurate and reliable segmentation of the land cover of a forest. Segmentation classes include forest areas, freshly deforested areas, deforested areas based on duration since deforestation, water bodies,

Introduction - Remote Sensing

thws Technische Hochschule Würzburg-Schweinfurt

Band number	Band description	Resolution (m)
B1	Coastal aerosol	60
B2	Blue	10
B3	Green	10
B4	Red	10
B5	Red-edge 1	20
B6	Red-edge 2	20
B7	Red-edge	20
B8	Near infrared (NIR)	10
B8A	Near infrared narrow	20
B9	Water vapour	60
B10	Shortwave infrared	60
B11	Shortwave infrared 1	20
B12	Shortwave infrared 2	20

Research Questions

1. How can weakly supervised/self-supervised semantic segmentation be effectively applied to multispectral satellite imagery with limited or weak annotation?
2. Do existing segmentation methodologies, generalize effectively for multispectral data, given its unique nature and variance from data used to train existing pre-trained models?
3. How does the spectral variability of multispectral satellite imagery, with its nine bands including RGB, impact the performance and generalization of existing segmentation models?

Source: Schmitt et al. (2022)

Related Works

Weakly Supervised Semantic Segmentation of Satellite Images for Land Cover Mapping- Challenges and Opportunities

This paper discussed the challenges of automatic land cover mapping from remote sensing imagery. The paper then introduces weakly supervised learning as a promising strategy, particularly when dealing with large-scale datasets with erroneous labels or at lower resolutions.

Motivation: the lack of a dedicated large training/benchmark dataset

Dataset

Sentinel 2 and sentinel 1 data

Training data: SEN12MS training subset (labelled using low resolution MODIS)

Validation data: SEN12MS hold-out subset

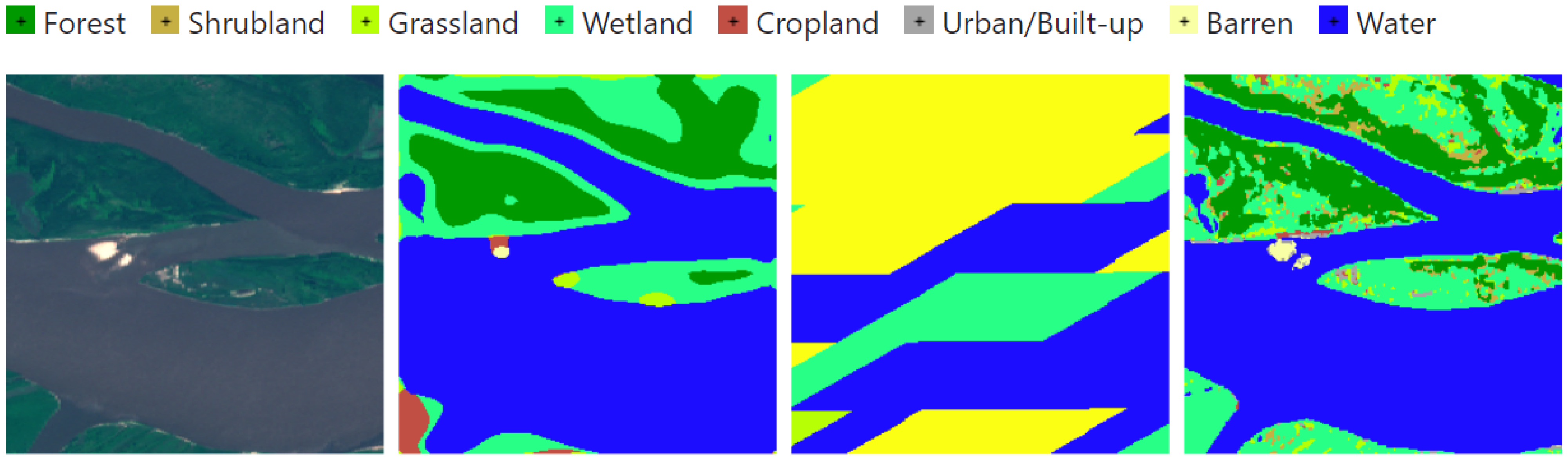
Test data: DFC2020 validation test (with high resolution labelling)

Training

U-net: Inputs: Number of channels and number of classes

Evaluation

Average accuracy, Confusion matrix



(a) Sentinel-2 (b) Prediction (c) SEN12MS label (d) DFC2020 label

Source: Schmitt et al. (2022)

Related Works

FreeSOLO: Learning to Segment Objects without Annotations

The paper introduces a fully unsupervised learning method that learns class-agnostic instance segmentation without any annotations.

FreeSOLO- Free Mask

1. Extract feature maps from images by a backbone model trained via self- supervision e.g. ResNet (or any CNN). The pre-trained model can be from supervised or unsupervised pre-training
2. Construct query **Q** and keys **K** from the features **I**. **I** is down-sampled bi-linearly to form the queries **Q**. **I** is used as the set of keys **K**
3. For each query in **Q** compute the cosine similarity with every key in **K** which results in **S**, the score maps
4. **S** is then normalised as soft masks by shifting the scores to the range [0, 1].
5. Compute the “maskness” score for each soft mask.
6. Convert soft masks to binary masks using a threshold τ
7. Binary masks are sorted by their “maskness” scores and redundant masks are removed via mask non-maximum-suppression (NMS). The resulting variable is the Free mask outputs.

Maskness score: a scoring function used to evaluate the quality of generated coarse masks. This scoring function, termed "maskness," is a non-parametric method calculated as

$$maskness = \frac{1}{N}\sum_{i=1}^{N} p_i$$

where ***N*** is the number of foreground pixels in the soft mask ***p***. Foreground pixels are those with values greater than a threshold ***τ***. Essentially, the maskness score gives higher weight to masks with confident foreground pixels and reduces the score for masks with uncertain foreground pixels.

Source:FreeSOLO Wang et al. (2022)

Next Steps

thws Technische Hochschule Würzburg-Schweinfurt

- Literature review
- Get data
- Data exploration