FH·W-S

Unsupervised Semantic Segmentation of Multispectral Satellite Imagery

Esther Ademola
5122738

March 2, 2024

Introduction

Climate change is an urgent global challenge, encompassing a spectrum of environmental transformations that affect ecosystems, weather patterns, and the delicate balance of the planet's climate. At the core of this intricate issue lies the interconnected relationship between human activities, serving as the primary catalyst for climate change [1]. Deforestation, the extensive removal of forests for diverse human endeavours, not only transforms landscapes but also plays a pivotal role in exacerbating the repercussions of climate change.

In addressing the imperative need to counteract the impacts of climate change and restore balance to our planet, a critical aspect is the monitoring and tracking of forest areas with regard to deforestation.This thesis focuses on the semantic segmentation of a forested area in Bavaria, Germany using advanced satellite remote sensing techniques. Specifically, the research leverages multispectral images of the Earth captured by the Sentinel-2 satellite at a high resolution of 10 meters by 10 meters and 20 meters by 20 meters (would be down-sampled for alignment) per pixel.

The segmentation includes categorising afforested areas (locations with vegetation), recently deforested areas, and older deforested areas based on the duration since deforestation, along with identifying water bodies and other significant features within the forest. It is also essential to segment features that are unknown/unnecessary.

Unfortunately, the majority of data in satellite monitoring lack proper labelling or have inaccuracies. Consequently, this task is approached as a weakly supervised/self-supervised/unsupervised semantic segmentation, utilising images with more spectral bandwidth than conventional RGB (Red, Green, and Blue) images. While existing segmentation methodologies, such as U-net, have demonstrated success in different datasets, scepticism arises regarding their optimization for our multispectral data, given its unique nature and considerable dataset variance from the data used to train existing pre-trained models.

Literature indicates a scarcity of research on unsupervised segmentation of multispectral images, particularly in the context of satellite imagery. This thesis aims to bridge this gap by developing a high-performance algorithm capable of unsupervised semantic segmentation, and subsequently weakly supervised classification to categorise different segments.

[1] https://www.ipcc.ch/2021/08/09/ar6-wg1-20210809-pr/

The significance of this research cannot be overstated, as it has the potential to mitigate deforestation and consequently reduce the adverse impacts of climate change on our environment. We have only one Earth, let's take care of it!

Aim

The primary aim of this thesis is to leverage advanced unsupervised semantic segmentation techniques on high-resolution multispectral data to achieve accurate and reliable segmentation of the land cover of a forest and then utilise weakly supervised classification to label the segments.

Related works

In recent years, several state-of-the-art architectures have been developed specifically for semantic segmentation, including U-Net[2], SegNet [1], and DeepLab. These models have been used with benchmark datasets for supervised segmentation of satellite images with great results [2]. Despite their effectiveness in supervised learning tasks, these models exhibit a limitation when applied to unannotated satellite data. Annotating large datasets is a costly and labour-intensive process, necessitating the development of models capable of performing weakly or unsupervised semantic segmentation.

Schmitt et al. proposes leveraging weakly supervised learning strategies, particularly utilising global land cover products with lower resolution and inherent noise. These products, while not ground truth, provide a substantial data source. The study utilises the SEN12MS dataset for training off-the-shelf models, such as Unet and Deeplab. Although the performance was not remarkable, the authors emphasised the potential for dedicated weakly supervised approaches to enhance high-resolution large-scale land cover mapping, as demonstrated by the identified limitations of existing off-the-shelf methods [3].

Likewise, in his thesis Regmi shows the application of the state-of-the-art unsupervised instance segmentation method, FreeSOLO [4], for satellite image segmentation. Overcoming the challenges of acquiring labelled data, the study benchmarks FreeSOLO on datasets such as iSAID, CrowdAI, and PASTIS, achieving significant results, including 0.9%AP50 on iSAID and 3.1%AP50 on CrowdAI. The method successfully segments diverse objects like buildings, water bodies, and roads. The research also emphasises the comparative performance of FreeSOLO-based weights in semantic segmentation downstream tasks, showcasing the advantages of self-supervised learning in scenarios with limited labelled data [5].

In summary, the reviewed literature underscores the effectiveness of state-of-the-art supervised semantic segmentation architectures. However, the challenge lies in their dependency on annotated data. This research aims to address this limitation by exploring models conducive to unsupervised or weakly supervised learning. By drawing insights from studies like [6], we aspire to contribute to the development of more versatile semantic segmentation models for multispectral images, extending their applicability to scenarios where labelled data is scarce or erroneous. This shift towards unsupervised learning holds significant potential for advancing semantic segmentation in remote sensing applications.

[2] `https://lmb.informatik.uni-freiburg.de/people/ronneber/u-net/`

Proposed Methodology

The segmentation of multispectral satellite images poses a unique set of challenges, primarily attributed to the intrinsic differences between these images and the datasets on which state-of-the-art pretrained models are typically trained—predominantly natural images with RGB channels. In response to these challenges, our proposed methodology comprises a multifaceted approach, combining the utilisation of existing pretrained models and the development of a bespoke unsupervised/self-supervised model.

To commence, we will explore the feasibility of employing existing pretrained models on our multispectral satellite image dataset. While recognising the significant advancements achieved by state-of-the-art pretrained models in various domains, it is imperative to acknowledge that these models were not originally trained on satellite images and, moreover, their training data predominantly consisted of RGB channels. Consequently, we approach this strategy with cautious optimism, understanding that the transferability of these models may be limited due to the disparate nature of the datasets.

Moving beyond the constraints of existing pretrained models, our second strategy involves the development of a customised model purpose-built for the unsupervised or self-supervised segmentation of multispectral satellite images. The rationale behind this approach is rooted in the need to bridge the gap between the unique characteristics of our dataset and the pre-existing models, which may not capture the intricacies of multispectral imagery. This tailored model, constructed from the scratch, will leverage self-supervised learning techniques.

Recognising the scarcity of labelled data, our third strategy entails the training of a weakly supervised classification model to enhance the accuracy of our segmentation results. Leveraging the limited labelled dataset at our disposal, this model will be designed to classify segments generated by the custom unsupervised model. By associating labels with predicted segments, this approach seeks to imbue the unsupervised model with enhanced contextual understanding, thereby refining the segmentation outcomes.

The evaluation of our proposed methodology will hinge on rigorous quantitative metrics, including Intersection over Union (IoU), Dice coefficient, and overall accuracy. These metrics will provide an objective basis for comparing the performance of the pretrained models and the custom unsupervised model. This comprehensive assessment will not only highlight the efficacy of each strategy but also guide the iterative refinement of the methodology.

In summation, our proposed methodology endeavors to provide a holistic and tailored approach to the unsupervised segmentation of multispectral satellite images. By combining the strengths of existing pretrained models, a bespoke unsupervised model, and weakly supervised learning techniques, we aim to address the unique challenges posed by multispectral data, ultimately paving the way for more accurate and contextually aware segmentation outcomes.

Proposed timeline

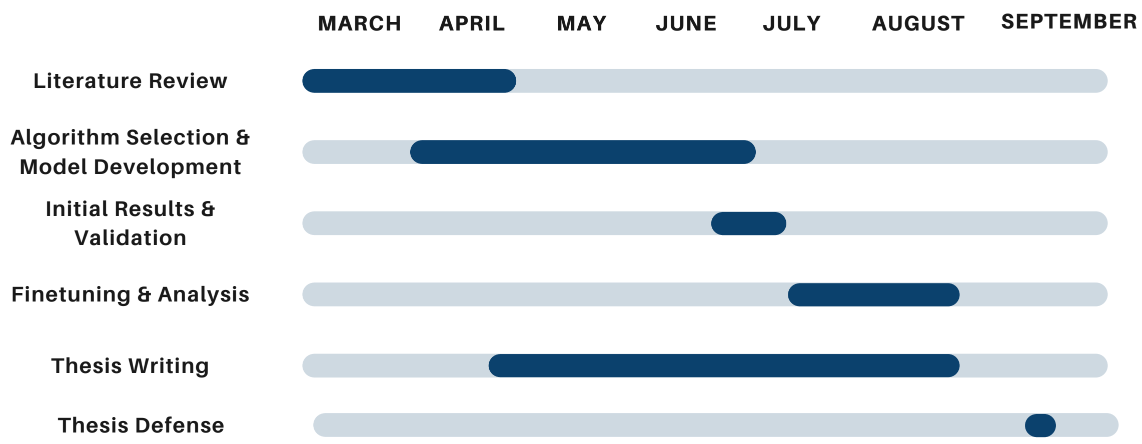

Figure 1: Proposed masters thesis timeline

References

[1] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," 2016.

[2] K. N. Ningthoujam Johny Singh, "Semantic segmentation of satellite images using deep-unet," *Arabian Journal for Science and Engineering*, 2022.

[3] M. Schmitt, J. Prexl, P. Ebel, L. Liebel, and X. X. Zhu, "Weakly supervised semantic segmentation of satellite images for land cover mapping – challenges and opportunities," 2020.

[4] X. Wang, Z. Yu, S. D. Mello, J. Kautz, A. Anandkumar, C. Shen, and J. M. Alvarez, "Freesolo: Learning to segment objects without annotations," 2022.

[5] S. Regmi, "Unsupervised image segmentation in satellite imagery using deep learning," 2023.

[6] A. Lassiter and M. Darbari, "Assessing alternative methods for unsupervised segmentation of urban vegetation in very high-resolution multispectral aerial imagery," *Plos One*, vol. 15, p. e0230856, 2020.