

Sion, November 28, 2023

**Project proposal:**

*Project topic - self-supervised pre-training for few-shot mapping drivers of tropical deforestation*

Tropical deforestation is a major environmental issue, resulting in carbon emissions, biodiversity loss, and impacting communities that rely on forests for food or shelter. To design effective policies to tackle this, it is necessary to know the drivers behind deforestation such as agriculture, mining, or fire. Efforts have been made toward the automatic recognition of drivers from satellite imagery using machine learning. However, training such models requires large amounts of annotated data which is difficult to obtain.

Recently, self-supervised learning (SSL) has shown great potential for reducing the amount of annotated data. SSL is a method of training models on pretext (auxiliary) tasks for which annotation is available, enabling the model to learn useful intermediate representations of input data. Common types of pretext tasks include contrastive (distinguishing between similar (positive) and different (negative) pairs of data points) and predictive (e.g., predicting a missing part of the input data), but a variety of other tasks have been proposed. Once trained on pretext tasks, the model can then be fine-tuned to the downstream task (in this case, classification of deforestation drivers) with significantly less annotated data required.

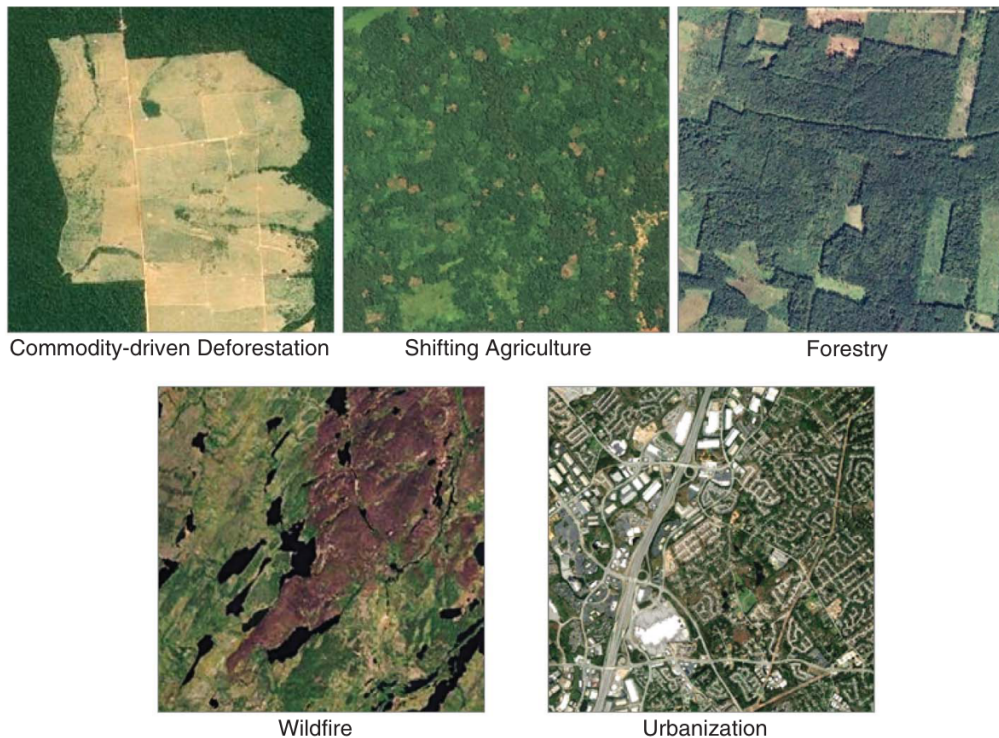


Figure 1. Examples of deforestation drivers (adapted from Curtis et al. 2018)

The aim of this project is to design and implement different pretext tasks, pre-train the model on these tasks, and evaluate its performance on the downstream task. In the first step, a baseline will be set using a model with randomly initialized weights or pre-trained on a common dataset (e.g., ImageNet). Then, one or multiple pretext tasks will be designed and the pre-training procedure will be defined and implemented. Finally, the pre-trained model(s) will be evaluated on the downstream task with varying amounts of annotated data to demonstrate the benefit of SSL.

The scope may be adapted to fit both a semester project or a master's thesis. In case you are interested or have questions, please feel free to reach out.

### **Bibliography**

- Irvin, J., Sheng, H., Ramachandran, N., Johnson-Yu, S., Zhou, S., Story, K., Rustowicz, R., Elsworth, C., Austin, K. and Ng, A.Y., 2020. Forestnet: Classifying drivers of deforestation in Indonesia using deep learning on satellite imagery. arXiv preprint arXiv:2011.05479.
- J. Pisl, L. H. Hughes, M. Rußwurm and D. Tuia, "Classification of Tropical Deforestation Drivers with Machine Learning and Satellite Image Time Series," IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium, Pasadena, CA, USA, 2023, pp. 911-914, doi: 10.1109/IGARSS52108.2023.10281472.
- Scheibenreif, L., Hanna, J., Mommert, M. and Borth, D., 2022. Self-supervised vision transformers for land-cover segmentation and classification. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 1422-1431).