Title: Energy-Efficient Crop Classification at Scale Using AgrUNet: A Multi-GPU Deep Learning Approach Within the MESEO Project

Monitoring agricultural parcels and identifying crop types are crucial pillars for sustainable soil management and agricultural resource optimization. With high-resolution remote sensing data from missions like Copernicus Sentinel-2 continuously expanding, there is an urgent need for automated, scalable, and computationally efficient tools capable of extracting phenological signatures over extensive geographic regions. To address this challenge, this work presents AgrUNet, a fast, scalable, and reliable Deep Learning model designed for pixel-level crop classification from multi-temporal and multi-spectral satellite image time series.

Developed within the framework of the EU Horizon MESEO project, AgrUNet is a Convolutional Neural Network based on the U-Net architecture. It utilizes 3D convolutions in its encoder and decoder branches to process spatial features, and integrates a multi-head temporal attention bottleneck, conditioned on a day-of-year positional encoding, to capture temporal growth dynamics. To meet the demands of Big Data Earth Observation (EO) and green computing, AgrUNet is built using PyTorch Lightning, enabling optimization for High-Performance Computing clusters through distributed multi-GPU processing and Automatic Mixed Precision (AMP). Within the MESEO software stack, the model is fully operationalized as a Docker-based microservice orchestrated by the GMV Prodigi® platform, seamlessly delivering automated, color-coded crop maps upon user requests.

AgrUNet was evaluated on two open-source datasets: one covering the Munich region (Germany) and another, OrthoPASTIS, covering four French regions. On the OrthoPASTIS, we achieve a peak multi-class Dice score of 0.86, ranging from 0.96 for meadows and 0.60 for sorghum, and similar performances are achieved for the Munich dataset. From a computational perspective, AgrUNet achieves a peak inference throughput of 419 img/s using AMP on a single NVIDIA GH200 node. Scaling to multiple nodes preserves near-linear speedup, reaching 838, 1226, and 1635 img/s on two, three, and four nodes, respectively. Compared to the U-TAE literature baseline, measured on the same hardware and time-series lengths, AgrUNet is 1.6× faster at matched precision and accuracy – rising to 2.0× with AMP – and 2.1× more energy-efficient.

This approach also brings clear sustainability benefits. Moving inference from CPU to GPU cuts energy consumption from 35.90 to 1.30 J per image, a reduction of roughly 96%, with a proportional decrease in carbon footprint. AgrUNet therefore shows that high-throughput, GPU-accelerated inference offers an accurate, scalable, and energy-efficient solution for operational EO applications, supporting land-use monitoring and agricultural sustainability.