trans4num INSPIRE HACKATHON 2024

FINAL REPORT ON CHALLENGE #3: AI-based cloud-free crop monitoring

Name of mentor(s): Petr Šimánek & Jiří Kvapil
Number of participants: 4

INTRODUCTION

Background of the Challenge:
Precision agriculture has emerged as a critical approach to enhancing crop yields, improving resource efficiency, and mitigating the negative impacts of environmental fluctuations. High-resolution satellite imagery (e.g., Sentinel-2) has proven invaluable for crop health monitoring and forecasting. However, persistent cloud cover, especially in regions with over 50% annual cloud coverage, significantly hampers consistent data acquisition. This leads to incomplete datasets and hampers timely interventions. Meanwhile, radar data (e.g., Sentinel-1) can penetrate cloud cover but may lack the spectral detail of optical imagery. Integrating diverse data sources to produce cloud-free, high-quality imagery and accurate vegetation indices such as NDVI remains a core challenge.

Full Explanation of the Challenge and Scope of Effort:
The overarching goal is to develop an AI-based spatiotemporal crop monitoring system that can effectively remove clouds from satellite images, thus enabling reliable crop health assessment and forecasting. The effort addresses two main objectives: (1) generating accurate, cloud-free, multi-spectral satellite imagery to track vegetation indices, and (2) forecasting crop dynamics under variable conditions, including cloud cover. We integrated optical (Sentinel-2) and radar (Sentinel-1) data, as well as supplementary contextual information like soil types and land surface analysis (LSA) products. By using state-of-the-art neural network architectures, including attention-based convolutional models and exploring diffusion-based generative approaches with autoencoders, our aim was to produce robust models that can handle incomplete and obscured data. The project's scope encompasses data preprocessing, model design, experimentation, evaluation, and the exploration of cutting-edge neural network techniques, including diffusion models

coupled with autoencoders, to refine cloud removal and improve NDVI forecasting accuracy.

METHODOLOGY

Team Description and Coordination:
The project team consists of researchers and engineers from the Faculty of Information Technology at the Czech Technical University in Prague and collaborators from Lesprojekt s.r.o. and HELP SERVICE – REMOTE SENSING s.r.o. The team coordinated closely with agricultural specialists and leveraged resources from European satellite data providers (Sentinel missions) to ensure data alignment and relevance. The cooperation extended to EUMETSAT for acquiring LSA products and to national mapping services (VUMOP) for soil-type information. Regular meetings were conducted to align goals, discuss data integration strategies, and validate the approaches against expert agricultural knowledge.

Technical Background:
Remote sensing techniques have evolved to provide multi-spectral satellite imagery critical for crop monitoring. However, frequent cloud cover reduces the availability of usable optical data. Researchers have proposed various algorithms and neural network architectures (e.g., U-Net, attention-based models, Generative Adversarial Networks) for cloud removal, and more recently, the UnCRtainTS architecture was introduced for robust cloud removal using multispectral time series. Furthermore, diffusion models have emerged as a cutting-edge generative framework capable of producing high-quality images from noise, and autoencoders provide effective dimensionality reduction and feature extraction. Integrating these approaches can potentially refine cloud-free imagery reconstruction and produce more accurate vegetation metrics.

Description of the Process of Solution:

1. **Data Preparation:**
    - **Satellite Imagery:** Sentinel-2 L2A frames (all 12 spectral bands) and Sentinel-1 GRD frames (VV and VH polarization) were acquired.
    - **Auxiliary Data:** LSA products (MDIDSSF, MDMETv3, MNSLF), Czech Republic soil-type maps from VUMOP, and agricultural land masks were integrated.
    - **Preprocessing:** Data were standardized to 10 m resolution, spatially aligned, and masked for clouds using Sentinel-2's Scene Classification Layer (SCL). Non-agricultural areas were optionally masked using VUMOP layers.

2. **Cloud Removal and Forecasting Model Development:**
    - **Baseline Models:**
        - Repeating the most recent cloudless Sentinel-2 frame
        - Mosaicking partially clouded frames based on cloud masks
        - A U-Net-based model as a convolutional baseline for cloud removal
    - **Advanced Architecture (UnCRtainTS):**
        - Implemented the UnCRtainTS state-of-the-art architecture for uncertainty-aware cloud removal.
        - Input sequences included the most recent Sentinel-2 and Sentinel-1 frames to facilitate accurate predictions under heavy cloud cover.
    - **Exploring Diffusion Neural Networks with Autoencoders:**
        - To further refine image quality and capture complex patterns under uncertain conditions, a diffusion-based generative framework was tested.
        - The approach integrated a diffusion model, which generates images by progressively denoising random noise into target images, with autoencoders that learned a lower-dimensional latent representation of cloud-free satellite scenes.
        - The diffusion model used guidance from the autoencoder's latent features to produce more coherent and stable outputs, potentially improving over direct prediction models in handling complex cloud patterns and subtle crop variations.
3. **Training and Optimization:**
    - Models were trained using MAE or MSE loss functions for pixel-wise reconstruction.
    - NDVI-specific training was performed by directly optimizing NDVI loss, focusing on accurately reconstructing vegetation indices.
    - Experiments varied input sequences (Sentinel-1 only, Sentinel-2 only, combined), integration of LSA products and soil maps, and batch sizes.
    - Hyperparameters (learning rates, batch sizes, loss functions) were tuned to balance computational efficiency and model accuracy.

Data & Equipment List:

- **Data Sources:** Sentinel-2 L2A (optical), Sentinel-1 GRD (radar), LSA products (EUMETSAT), soil-type data (VUMOP), and agricultural masks.
- **Equipment and Tools:** NVIDIA GPU clusters for training deep learning models, Python libraries (PyTorch, TensorFlow), geospatial data processing tools (GDAL), and cloud storage for data handling.

- **Software Environment:** Linux-based HPC clusters, Docker containers for reproducible environments, Git repositories for version control.

Detailed Implementation Plan:

- **Step 1: Dataset Assembly:** Aggregate Sentinel-1, Sentinel-2, and auxiliary data into a single coherent dataset.
- **Step 2: Preprocessing and Masking:** Apply SCL-based cloud masks, and align all data to a uniform 10 m grid.
- **Step 3: Model Development:**
  - Implement baseline U-Net and UnCRtainTS architectures.
  - Integrate diffusion models and autoencoders into the pipeline to refine image synthesis.
- **Step 4: Training and Validation:** Train models using year 2022 data for training and 2023 data for validation, focusing on April to October.
- **Step 5: Analysis and Refinement:** Evaluate MAE, MSE, and NDVI metrics; incorporate LSA and soil-type data; refine model hyperparameters and architectures based on performance.
- **Step 6: Testing Diffusion + Autoencoders:** Compare the diffusion-based approach with baseline UnCRtainTS predictions for visual quality and NDVI accuracy, iterating as needed.

Analysis of Needs of Stakeholder Groups:

- **Farmers and Agronomists:** Require timely and accurate NDVI maps to plan irrigation, fertilization, and pest control measures.
- **Researchers and Extension Services:** Need high-quality, cloud-free datasets to analyze trends, develop decision support systems, and ensure long-term food security.
- **Policy Makers and Environmental Agencies:** Benefit from accurate data on vegetation health and land use for informed resource allocation, environmental assessments, and sustainability planning.

Experimental Results:

- Models integrating Sentinel-1 and Sentinel-2 data significantly outperformed single-modality approaches.
- UnCRtainTS architecture showed improvement over U-Net and baseline methods in terms of MAE and MSE metrics.
- Training on agricultural pixels and optimizing NDVI metrics directly improved vegetation index prediction.
- Auxiliary data (soil types, LSA products) had a limited impact on performance in initial trials.

- Preliminary tests with diffusion models and autoencoders indicated the potential for smoother and more detailed reconstructions, though further hyperparameter tuning and computational resources are needed.

FINDINGS & CONCLUSION

Discussion of the Results and Findings:
The integrated approach—combining Sentinel-1 and Sentinel-2 imagery—delivered more reliable predictions under cloud cover. UnCRtainTS outperformed traditional U-Net baselines, indicating that attention-based temporal aggregation and uncertainty modeling are beneficial. Although additional data (soil types, LSA) did not yield significant improvements, their value may emerge with more specialized feature engineering or in different use cases. Direct NDVI optimization proved effective in improving the quality of vegetation index forecasts.

The exploration of diffusion models combined with autoencoders shows promise. While initial experiments did not immediately surpass UnCRtainTS in standard metrics, the generated imagery appeared structurally consistent, suggesting potential advantages once training and architectural refinements are more thoroughly explored. These generative approaches may offer new avenues for handling extreme cloud conditions and complex environmental patterns.

Further Improvements:

- **Refinement of Diffusion-Based Methods:** Fine-tuning the diffusion model's parameters, improving autoencoder architectures, and increasing the training dataset volume could yield substantial gains in image quality and NDVI accuracy.
- **Enhanced Data Integration:** Incorporating reliable point data (e.g., temperature, precipitation, wind) or refining the use of soil and LSA data with advanced feature extraction methods may further enhance the model's robustness and predictive capability.
- **Scalability and Generalization:** Scaling the approach to larger regions, different climatic zones, and diverse crop types can be pursued. Additional validation against ground truth measurements and domain expert feedback will help ensure adaptability and usability across agricultural contexts.
- **Operationalizing the Models:** Streamlined inference pipelines and user-friendly interfaces can facilitate deployment for agronomists, policymakers, and other stakeholders. Overcoming computational constraints (e.g., by implementing model compression or using cloud-based inference) will make the system more accessible and practical.

In conclusion, the integration of advanced neural network architectures, including UnCRtainTS and exploratory diffusion-plus-autoencoder methods, presents a promising path forward in generating cloud-free satellite imagery and accurate crop health forecasts. This paves the way for more efficient and sustainable precision agriculture practices, ensuring better resource management, timely interventions, and ultimately, improved food security.