Toward Plant Disease Detection with EnMAP Hyperspectral Imagery

A Proposed CNN–ViT Hybrid Deep Learning Framework

Asherl Bwatiramba[1,2] **Irina Zlotnikova**[1] **Rajalakshmi Selvaraj**[1]
Dimane Mpoeleng[1]

[1] Botswana International University of Science and Technology (BIUST), Botswana
[2] Botho University, Botswana

African Astronomical Society Conference, Kasane, Botswana, March 27, 2026

What This Presentation Is - And Is Not

What this presentation is:

- A **problem framing** for plant disease detection from spaceborne hyperspectral imagery
- A **method proposal** explaining why a CNN–ViT hybrid is worth testing
- A **benchmark design** for comparing CNN, ViT, and hybrid models fairly
- a **study roadmap** for future empirical validation

What it is not:

- Not a completed paper with results
- Not a final accuracy comparison
- Not a field-validated disease map
- Not yet a transferability study

Scope statement. This talk does *not* present final disease-detection results. It presents the research gap, the proposed architecture, and the evaluation plan that will be used in the next phase of the study

Why this presentation still matters? Our contribution is methodological: it sets up a credible route for testing whether hybrid deep learning can improve disease-sensitive analysis of EnMAP imagery under realistic label constraints

Why This Problem Remains Open

Opportunity from hyperspectral remote sensing:

- Narrow bands can capture stress responses linked to chlorophyll, water, and biochemical change
- Spaceborne hyperspectral imaging expands disease monitoring beyond plot- or UAV-scale studies
- EnMAP offers richer spectral detail than broad-band multispectral missions

Why the solution is not straightforward

- **Satellite multispectral data:** scalable but often too spectrally coarse
- **UAV hyperspectral data:** detailed but spatially limited
- **Pure CNNs:** strong local extraction but weak long-rang reasoning
- **Pure ViTs:** strong global context but typically more data hungry

Identified methodological gap

Approach	Local	Global
CNN	High	Low
ViT	Mid	High
CNN–ViT	High	High

Local refers to fine-scale nearby features in the image patch, while *Global* refers to broader whole-patch context and long-range relationships.

Working proposition. A hybrid architecture may be better suited to agricultural hyperspectral imagery because it can combine:

- local spectral-spatial cues
- long-range inter-band dependencies
- potentially better behaviour under limited labels

Proposed Framework and Workflow

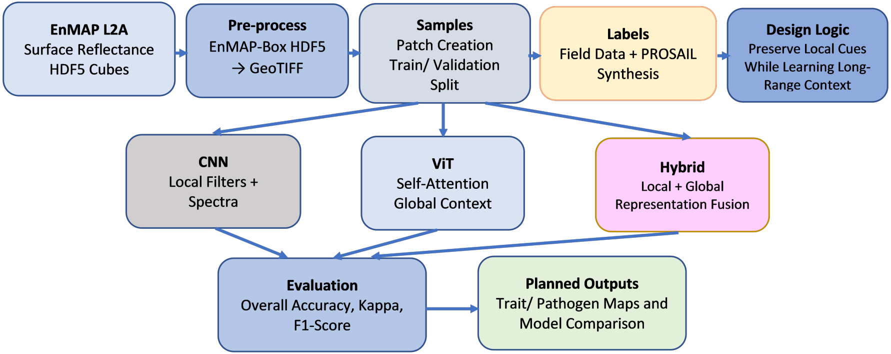

Figure 1. Workflow of the proposed CNN–ViT hybrid framework for plant disease detection from EnMAP hyperspectral imagery

Planned Benchmark and Study Setting

Study area and data plan

- **Location:** Pandamatenga, Northeastern Botswana
- **Extent:** about 450 km^2 of agricultural land
- **Input:** EnMAP Level-2A surface reflectance
- **Spectral basis:** 224 bands across visible, NIR, and SWIR
- **Environment:** EnMAP-Box in QGIS for impor conversion, and experiment setup

Benchmark logic

- Identical preprocessing across all models
- Identical train / validation splits
- OA, Kappa, and F1 for direct comparison

What the benchmark is intended to test

Model	Expected Strength(s)	Main Risk(s)
CNN	Efficient local extraction from spectral-spatial patches	Misses long-range context
ViT	Broad contextual relations and inter-band dependencies	Data hungry; may overfit
Hybrid	Combines local inductive bias with global attention	More complex pipeline

Note:

- These are **planned comparisons**, not completed results
- The benchmark is proposed here so that later results can be interpreted credibly
- The aim is transparent model comparison, not premature performance claims

What Is Still Missing and What Comes Next

Included in this talk

- research motivation and gap
- proposed CNN–ViT framewo
- workflow and benchmark design
- intended scientific value

Not yet completed:

- No final disease-detection results
- No completed CNN vs. ViT v hybrid comparison
- No large-scale field-validated label set
- No transferability test across seasons or sites

Immediate next steps:

- Strengthen field-labelled validation
- Run the full benchmark
- Test generalisation across seasons and locations
- Extend toward multi-mission hyperspectral learning

Conclusion. This is a **proposed approach presentation**. It defines a path for evaluating whether a CNN–ViT hybrid can improve plant disease detection from EnMAP hyperspectral imagery under realistic data and label constraints