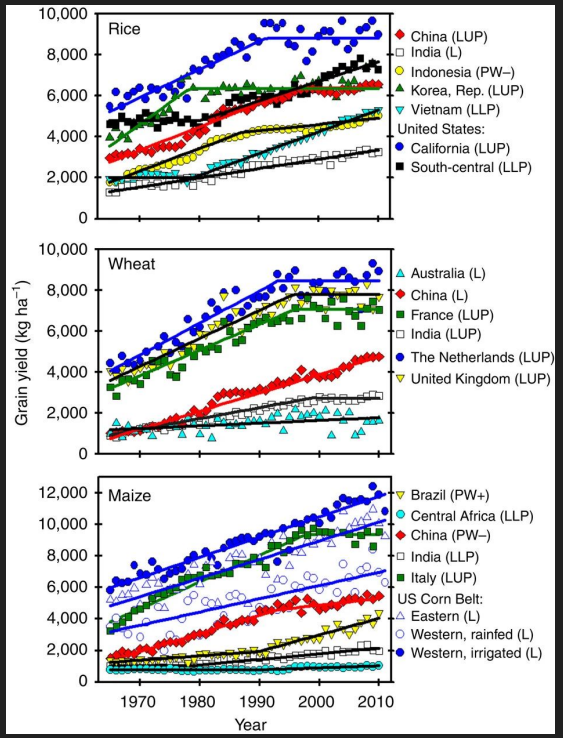
PhytoOracle: scalable, modular phenomic data processing pipelines

Emmanuel Gonzalez, Ariyan Zarei, Nathanial Hendler, Michele Cosi, Jeffrey Demieville, Travis Simmons, Holly Ellingson, Nirav Merchant, Eric Lyons, Duke Pauli

Yield improvements from historical innovations have been depleted

(American Society for Engineering Education)



(Grassini et al., 2013)

Emerging technologies are part of the solution

Proximal sensors generate large datasets which require complex cyberinfrastructure

Robots

(Wall Street Journal, LemnaTec)

Carts

(USDA)

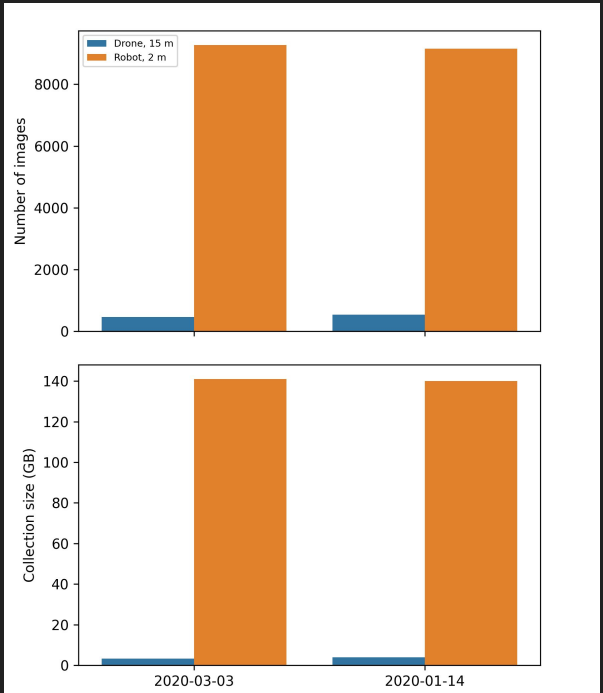
Drones

(DJI)

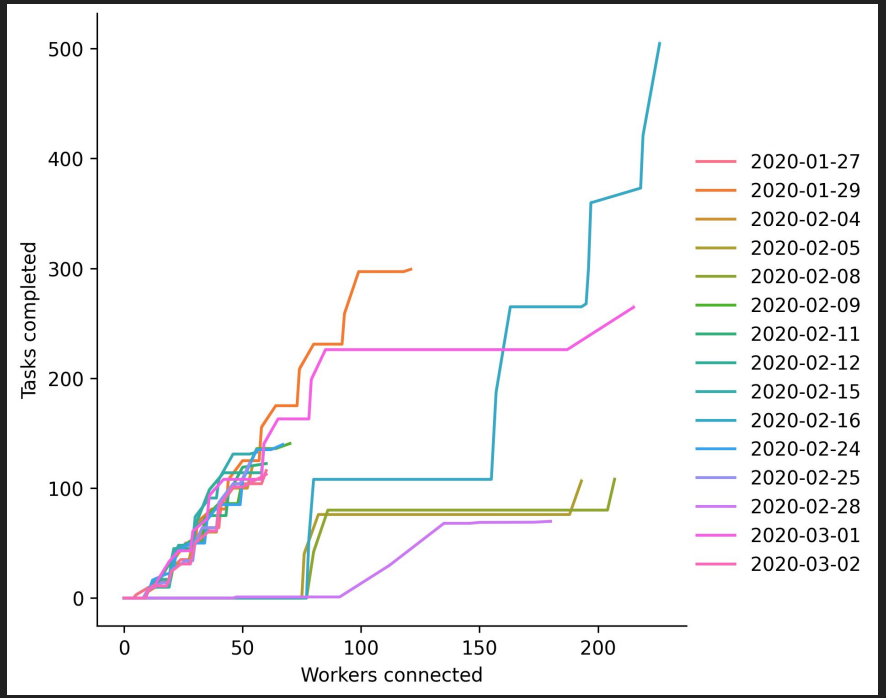
Phones

(IITA)

New technologies bring new bottlenecks



Increasing data size



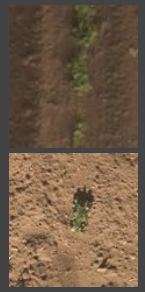
Need for scalable frameworks

PhytoOracle provides a general-use framework

GeoTIFF image/s + GeoJSON/Shapefile + PyTorch Faster RCNN model

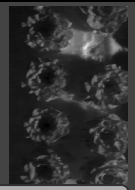
RGB

Faster R-CNN



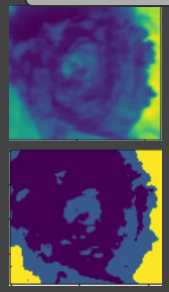
Bounding area

Thermal



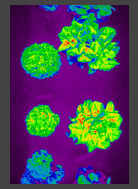
Faster R-CNN

K-means

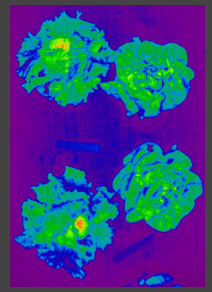


Canopy temperature

PSII

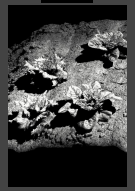


FLIP

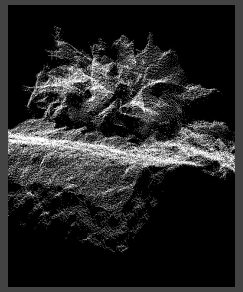


Photochemical efficiency

3D

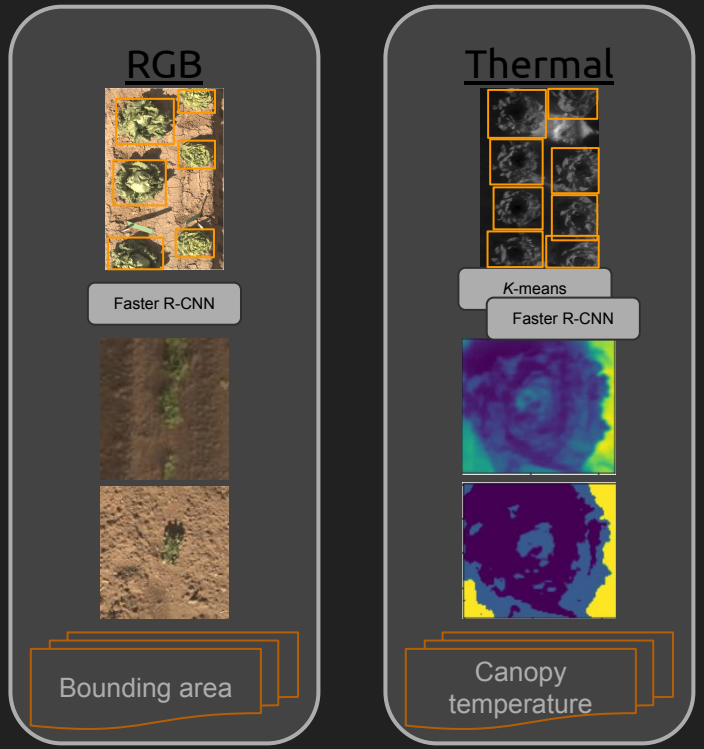


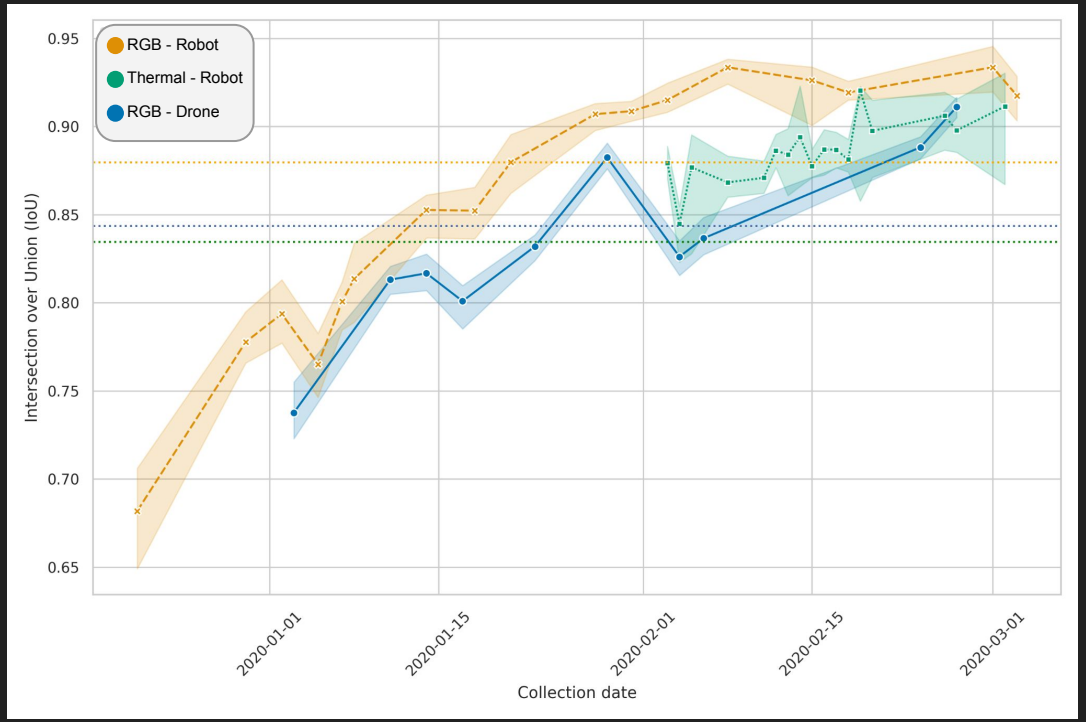
Semantic segmentation



Aboveground biomass

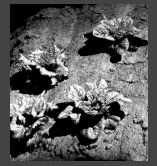
PhytoOracle provides a general-use framework





PhytoOracle enables reproducible science

3D



Pre-processing

Sequential alignment

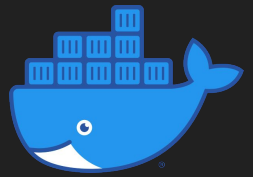
Post-processing

Plant clip

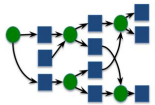
Aboveground biomass

```
FROM ubuntu:18.04

WORKDIR /opt
COPY . /opt

USER root

ARG DEBIAN_FRONTEND=noninteractive
RUN apt-get -o Acquire::Check-Valid-Until=false -o Acquire::Check-Date=false update -y
RUN apt-get install -y python3.6-dev \
                       python3-pip \
                       wget \
                       gdal-bin \
                       libgdal-dev \
                       libspatialindex-dev \
                       build-essential \
                       software-properties-common \
                       apt-utils \
                       libsm6 \
                       libxext6 \
                       libxrender-dev \
                       libgl1-mesa-dev

RUN add-apt-repository ppa:ubuntugis/ubuntugis-unstable
RUN apt-get update --fix-missing
RUN apt-get install -y --fix-missing libgdal-dev
RUN pip3 install cython
RUN pip3 install --upgrade cython

RUN pip3 install pyproj==1.9.6
RUN pip3 install numpy==1.19.1
RUN pip3 install opencv-python==3.4.2.16
RUN pip3 install opencv-contrib-python==3.4.2.16
RUN pip3 install open3d==0.11.2

ENTRYPOINT [ "/usr/bin/python3", "/opt/main.py" ]
```

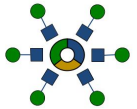
Tools



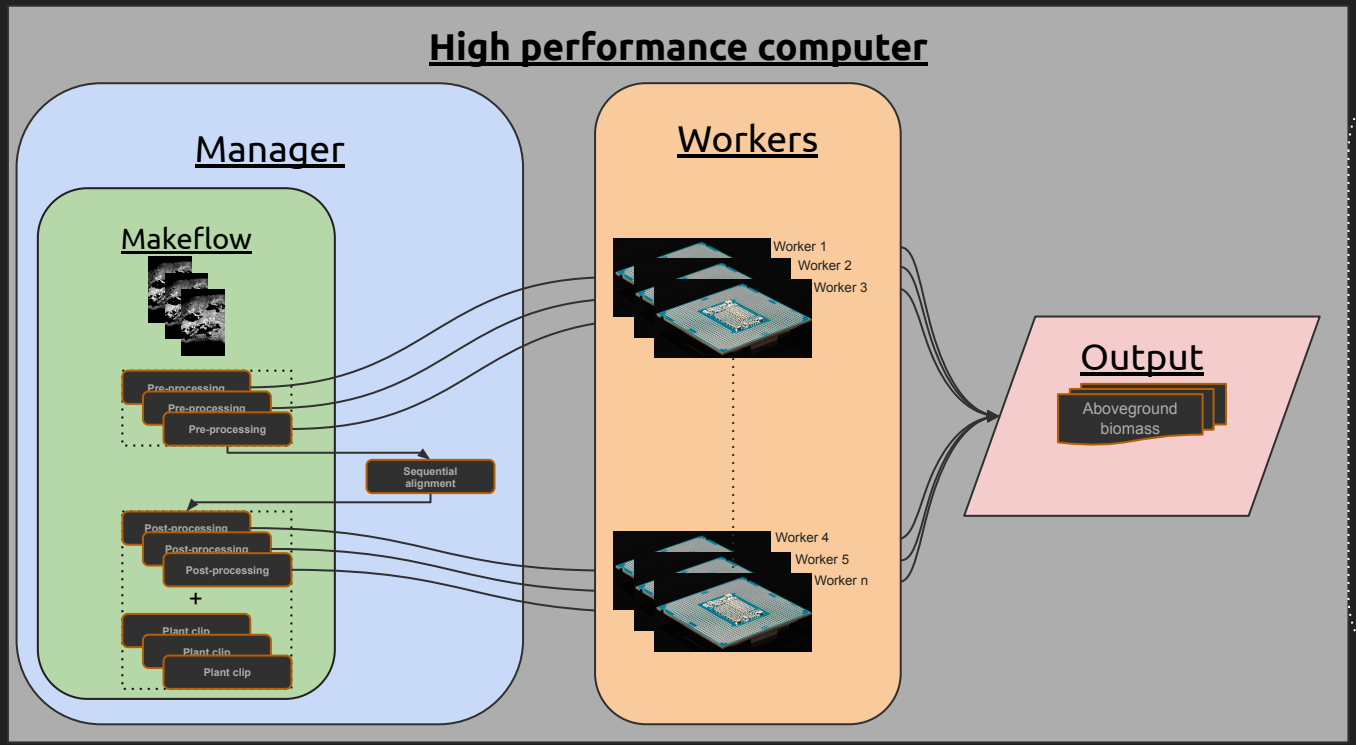
Makeflow



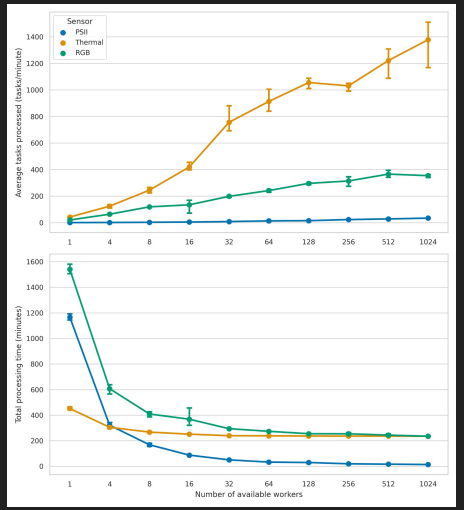
Work Queue



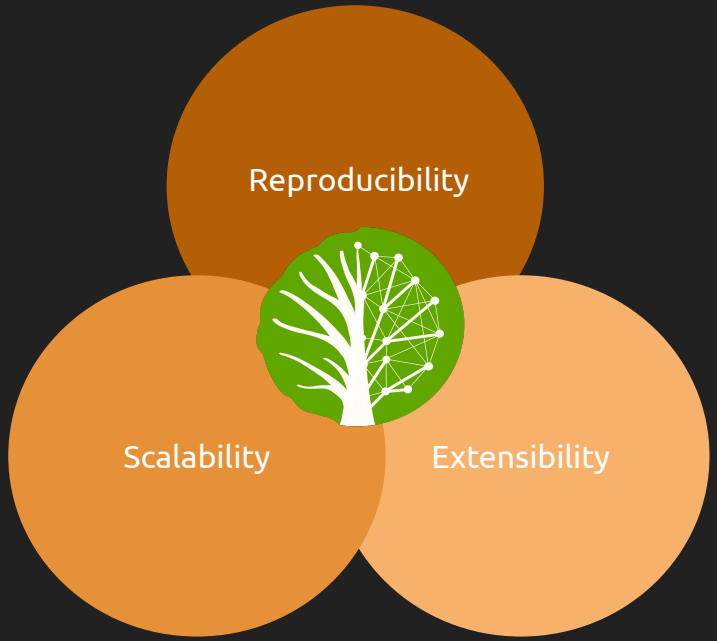
PhytoOracle leverages scalable processing



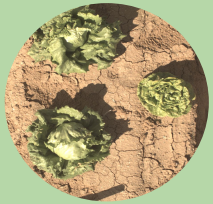
Benchmarking



PhytoOracle addresses three major concerns

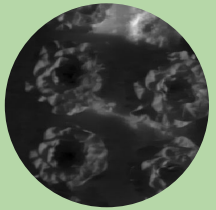


RGB



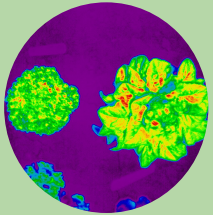
Disease detection
Canopy area

Thermal



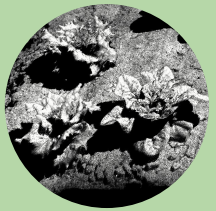
Water use efficiency
Stomatal conductance

Fluorescence



Photosynthetic efficiency
Stress assessment

3D / Lidar



Architecture
Biomass
Leaf angle

Acknowledgements

The University of Arizona

U.S. Department of Energy

CyVerse
Transforming SCIENCE through DATA-DRIVEN DISCOVERY

arpa·e

Saint Louis University

The George Washington University
Washington, DC

Donald Danforth Plant Science Center

The University of Arizona
High Performance Computing

Work Queue

Makeflow

docker