

TREE-D: Tree Recognition, Evaluation, and Extraction Dataset

Sean Keane¹, Lucas Kinder¹, Paul Kinder¹, Gregory Dahle², Aaron Maxwell³

¹Natural Resource Analysis Center, West Virginia University, Morgantown, WV

²Division of Forestry, West Virginia University, Morgantown, WV

³Department of Geology and Geography, West Virginia University, Morgantown, WV



Abstract

This project aims to establish a comprehensive, open-source benchmark dataset of tree canopies captured via unmanned aerial systems (UAS) using RGB, LiDAR, and multispectral imagery. West Virginia University's Natural Resource Analysis Center is developing standardized protocols for data collection and annotation, ensuring precise delineation of tree crowns with bounding boxes and polygons. This resource will accelerate the development of machine learning models for automated tree species identification. The dataset includes evaluation criteria for contributors, guidelines for standardized collection across seasons and locations, and protocols for data management. This initiative will serve researchers and land managers by reducing resources needed for model development while increasing accuracy across diverse environments.

Standardized Data Format

```
{info:
  description: "UAV Tree Crown Detection Dataset"
  url: "https://github.com/forest-lab/uvtdc"
  version: "1.0"
  year: "2025"
  contributor: "Sean Keane"
  date_created: "2025-01-23"
}
```

```
{licenses:
  id: 1
  name: "MIT License"
  url: "https://opensource.org/licenses/MIT"
}
```

```
{categories:
  id: 1
  species: "Quercus rubra"
  genus: "Quercus"
  family: "Fagaceae"
}
```

```
{images:
  id: 1
  file_name: "IMG_001.tif"
  width: 1024
  height: 1024
  date_captured: "2024-08-06"
  julian_day: 219
  time_captured: "16:30"
  license: 1
  sensor: "Sentinel2"
  altitude: 60
  focal_length: 8
  sensor_width: 7.0656
  sensor_height: 5.2992
  resolution: 0.02
  spectral_bands: {5 keys}
  state: "West Virginia"
  county: "Monongalia"
  location_description: "Little Indian Creek WMA"
}
```

```
{(Untitled-1):
  info: {6 keys}
  licenses: [1 items]
  categories: [1 items]
  images: [1 items]
  annotations: [1 items]
}
```

Data Includes:

- Julian day
- Sensor specs
- Spectral band information
- Flight parameters
- Location description
- Taxonomical classification

```
{spectral_bands:
  blue: {3 keys}
  green: {3 keys}
  red: {3 keys}
  redEdge: {3 keys}
  nir: {3 keys}
}
```

```
{blue:
  wavelength: 475
  bandwidth: 30
  order: 1
}
```

```
{green:
  wavelength: 555
  bandwidth: 20
  order: 2
}
```

```
{red:
  wavelength: 670
  bandwidth: 30
  order: 3
}
```

```
{redEdge:
  wavelength: 715
  bandwidth: 10
  order: 4
}
```

```
{nir:
  wavelength: 840
  bandwidth: 20
  order: 5
}
```

```
{segmentat...
```

```
120
```

```
{segmentat...
```

```
100
```

```
{segmentat...
```

```
140
```

```
{segmentat...
```

```
110
```

```
{segmentat...
```

```
160
```

```
{bbox:
  110
```

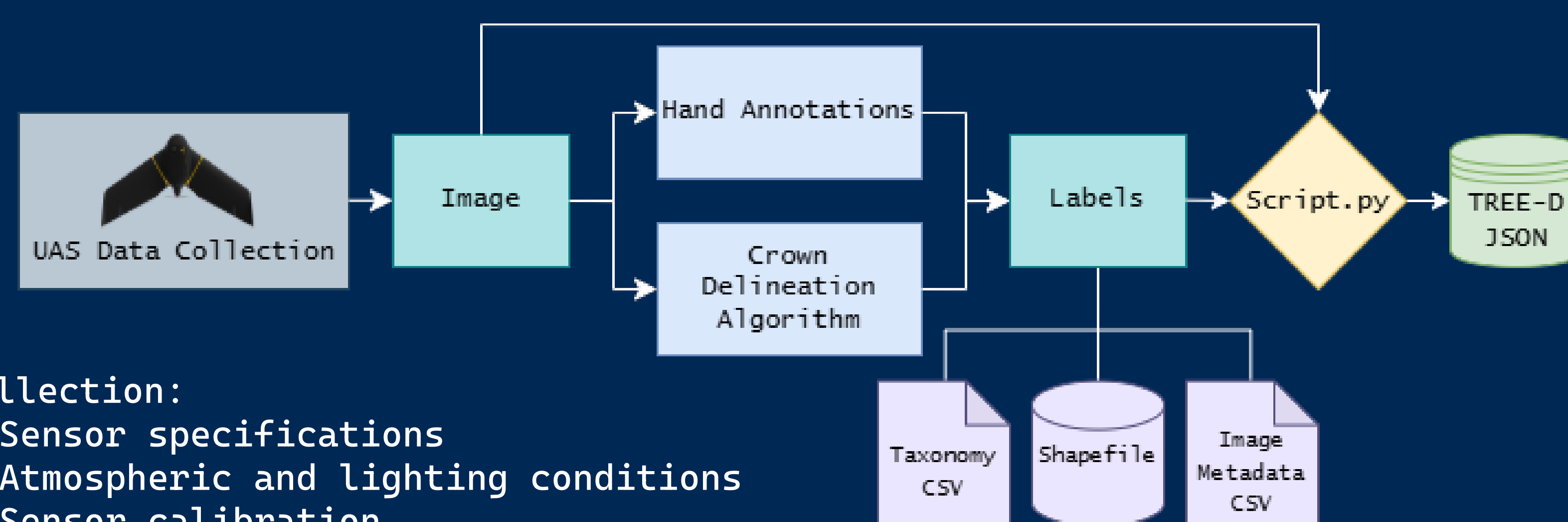
```
{bbox:
  100
```

```
{bbox:
  50
```

```
{bbox:
  50
```

```
{annotations:
  id: 1
  image_id: 1
  category_id: 1
  segmentation: [5 items]
  area: 800
  bbox: [4 items]
  iscrowd: 0
}
```

Data Integrity & Workflow



Collection:

- Sensor specifications
- Atmospheric and lighting conditions
- Sensor calibration

Processing:

- Radiometric corrections
- Photogrammetry parameters
- LiDAR noise reduction

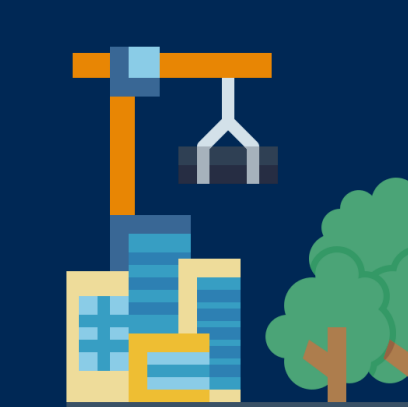
Post-Processing:

- Image artifact removal
- LiDAR derivative quality
- Annotation procedures and efficacy



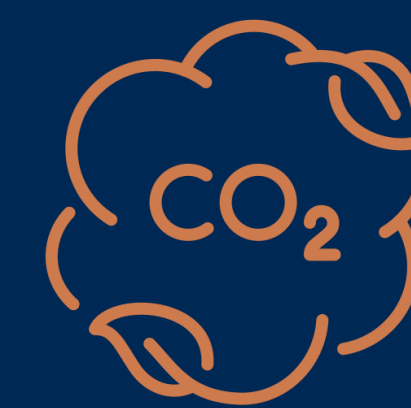
Potential Uses & Future Projects

Tree species identification and vegetation mapping



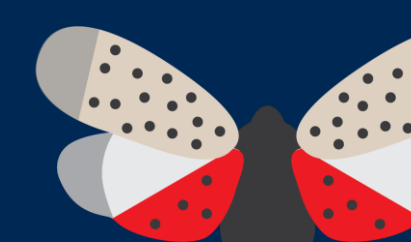
Urban forest inventory and diversity assessments

Carbon sequestration and biomass estimation models



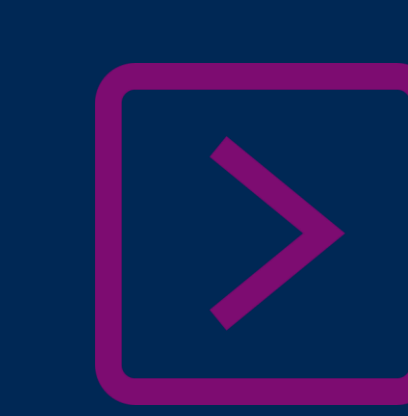
Post-storm damage assessment and recovery planning

Invasive species detection and monitoring programs



Spectral feature selection and data engineering

Seasonal phenology studies across different ecosystems



Open-source, commercial, and academic research

What's Next?

- Expand dataset diversity across geographic regions and seasons
- Streamline contribution validation with quality assessment tools
- Establish partnerships with government agencies, universities, and municipalities
- Create tools for visualization and custom dataset generation
- Benchmark machine learning models to establish performance baselines
- Release training and annotation tutorials

Acknowledgments

Funding for this project has been provided through the Tree Research & Education Endowment Fund and the Utility Arborist Research Fund



Learn More & Contribute!

We welcome contributions from researchers, arborists, remote sensing specialists, or anyone with data they can share!

We hope that this project will provide valuable data to future projects and accelerate research and collaboration!

If you would like to contribute, please contact LNKINDER@MAIL.WVU.EDU or stop by the Natural Resource Analysis Center (AgScience 4203).



[github/smk0061/TREE-D](https://github.com/smk0061/TREE-D)