Rooftop Solar Photovoltaic Detection and Verification System

1. Model Overview

- **Model Name:** Rooftop Solar PV Detection and Verification Pipeline
- **Model Type:** Deep learning–based segmentation and geospatial verification system
- **Base Architecture:** YOLOv8m-seg (Ultralytics)
- **Pretraining:** Pretrained on the COCO dataset for generic object detection and segmentation
- **Fine-tuning:** Fine-tuned on multiple rooftop solar panel datasets from Roboflow Universe
- **Primary Task:** Detection, segmentation, and verification of rooftop solar photovoltaic installations using high-resolution satellite imagery

2. Datasets Used

The model was fine-tuned using curated subsets of three publicly available rooftop solar datasets from Roboflow Universe. Only selected samples were used—not the full datasets.

Dataset 1: Custom Workflow 2 – Instance Segmentation Dataset

Used for: Dense rooftop PV segmentation patterns and varied roof textures

```
@misc{custom-workflow-2-instance-segmentation-di56e_dataset,
  title        = {Custom Workflow 2 Instance Segmentation Dataset},
  type         = {Open Source Dataset},
  author       = {test},
  howpublished = {\url{https://universe.roboflow.com/test-2bhme/custom-workflow-2-instance-
segmentation-di56e}},
  url          = {https://universe.roboflow.com/test-2bhme/custom-workflow-2-instance-
segmentation-di56e},
  journal      = {Roboflow Universe},
  publisher    = {Roboflow},
  year         = {2025},
  month        = {jun},
  note         = {visited on 2026-01-04}
}
```

Dataset 2: Solar Panel Detector Dataset

Used for: General rooftop PV appearance and negative rooftop examples

```
@misc{solar-panel-detector-imvoh_dataset,
  title        = {Solar-Panel-Detector Dataset},
  type         = {Open Source Dataset},
  author       = {Ariel Drabkin},
  howpublished = {\url{https://universe.roboflow.com/ariel-drabkin-tifqg/solar-panel-detector-
imvoh}},
  url          = {https://universe.roboflow.com/ariel-drabkin-tifqg/solar-panel-detector-
imvoh},
```

```
  journal      = {Roboflow Universe},
  publisher    = {Roboflow},
  year         = {2023},
  month        = {dec},
  note         = {visited on 2026-01-04}
}
```

Dataset 3: GEO-Solar Dataset

Used for: Geographically diverse rooftop layouts and panel orientations

```
@misc{geo-solar-tooai_dataset,
  title        = {GEO-Solar Dataset},
  type         = {Open Source Dataset},
  author       = {GEOproject},
  howpublished = {\url{https://universe.roboflow.com/geoproject/geo-solar-tooai}},
  url          = {https://universe.roboflow.com/geoproject/geo-solar-tooai},
  journal      = {Roboflow Universe},
  publisher    = {Roboflow},
  year         = {2023},
  month        = {may},
  note         = {visited on 2026-01-04}
}
```

Dataset Characteristics (Common)

- **Format:** YOLOv8 segmentation format
- **Source:** Roboflow Universe
- **Content:** High-resolution satellite imagery with rooftop solar annotations
- RGB satellite imagery
- Pixel-level masks
- Rooftop-mounted solar photovoltaic panels
- **Variations in:** Roof material, panel orientation, illumination and shadow, urban and semi-urban environments

3. Data Preprocessing

- Images resized to 640 x 640 pixels
- Normalization handled internally by the YOLOv8 pipeline
- No synthetic data augmentation applied
- Original training/validation splits preserved from dataset exports

4. Model Architecture

- **Backbone:** Convolutional feature extractor pretrained on COCO
- **Neck:** Multi-scale feature aggregation for spatial robustness

- **Head:** YOLOv8 segmentation head producing:
  - Pixel-level masks
  - Confidence scores per detection

5. Inference and Decision Logic

The trained YOLOv8 model operates within a multi-stage verification pipeline:

1. Satellite image retrieval (Google Static Maps with ESRI fallback)
2. Image quality assessment
3. Segmentation inference
4. Mask-to-polygon conversion
5. Rooftop buffer intersection validation (1200 / 2400 sqft)
6. Solar area and capacity estimation
7. Euclidean distance calculation (centroid to query point)
8. Confidence calibration
9. Final quality control classification

This design ensures interpretable, auditable predictions rather than raw detections.

Euler Distance Verification Metric

Purpose: Validate spatial correctness of detected PV panels

$$dx_m = (x_{centroid} - x_{center}) \times mpp$$

$$dy_m = (y_{centroid} - y_{center}) \times mpp$$

$$distance = \sqrt{(dx_m^2 + dy_m^2)}$$

Interpretation:

- 0–5 m → Excellent localization
- 5–15 m → Acceptable (geocoding jitter)
- 15+ m → Potential mis-association

6. Assumptions

- Solar panels are visually distinguishable in high-resolution satellite imagery
- Average panel area ≈ 1.9 m²
- Power density ≈ 190 W/m²
- Coordinate localization error within ±10 meters
- Rooftop solar arrays exhibit structured geometric patterns

7. Known Limitations

- Reduced performance under heavy cloud cover
- Shadows may partially obscure panels
- Very small rooftop installations may be missed
- Reflective rooftops can cause false positives
- Dependent on satellite image resolution and recency

8. Bias and Fairness Considerations

- Potential geographic bias due to dataset distribution
- Higher accuracy in dense urban regions
- Reduced robustness for uncommon roof materials

Mitigation Strategies:

- Conservative confidence thresholds
- Explicit NOT_VERIFIABLE classification under uncertainty

9. Failure Modes

The system explicitly handles:

- Image acquisition failures
- Poor image quality
- Inference errors
- Ambiguous rooftop structures

All such cases are labeled NOT_VERIFIABLE with descriptive reason codes.

10. Retraining Guidance

Retraining is recommended when:

- Deploying to new geographic regions
- Using different satellite imagery providers
- Encountering new rooftop construction styles

Recommended Steps:

- Add region-specific annotated data
- Revalidate buffer thresholds
- Recalibrate confidence scoring

11. Ethical Considerations

- No personally identifiable information is processed
- Uses only publicly available satellite imagery

- Outputs are transparent, auditable, and explainable

12. Intended Use

Intended for:

- Academic research
- Sustainability assessment
- Energy planning and evaluation

Not intended for:

- Legal enforcement
- Individual household monitoring

13. License

Developed for academic, research, and evaluation purposes.