Surya Saathi: Model Card

AI-Powered Solar Subsidy Verification System

---

1. Model Overview

Model Name: YOLOv11 Solar Panel Detection Model
Version: 1.0.0 | **Release Date:** December 2025
Framework: Ultralytics YOLOv11 | **Model Size:** 5 MB
Owner: AlgoMinds Team

Purpose

Surya Saathi automates fraud detection in India's rooftop solar subsidy program by verifying actual solar panel installations using satellite imagery analysis and multi-stage verification pipeline.

Key Components

1. **Photo Forensics:** EXIF/GPS metadata validation
2. **Satellite Detection:** YOLOv11 panel detection on pre/post-installation imagery
3. **Equipment Verification:** OCR-based ALMM database validation
4. **Shadow Analysis:** Temporal/spatial validation using pvlib
5. **Decision Engine:** Weighted ensemble scoring for auto-approval/manual review/rejection

---

2. Technical Specifications

Specification	Details
Input	Satellite/Aerial RGB images (640×640 to 1280×1280px)
Output	Bounding boxes, confidence scores, panel count estimates, QC status
Architecture	YOLOv11n (nano, ~5MB)
Training Data	5,000 satellite image pairs (70% train, 15% val, 15% test)
Training Time	2.8 hours (50 epochs, NVIDIA RTX 4090)
Inference Speed	45-65ms/image (GPU), 180ms/image (CPU)
Throughput	15-22 images/sec (GPU)

Performance Metrics (Test Set)

Metric	Score	Notes
mAP@0.5	**0.824**	Overall detection accuracy
Precision	0.765	True positives among predictions
Recall	0.787	True positives among ground truth
F1-Score	0.776	Harmonic mean
Confidence Distribution	Mean: 0.782, Std: 0.142	Prediction confidence

Hyperparameters

- **Optimizer:** SGD (momentum=0.937)
- **Learning Rate:** 0.001 (cosine annealing)
- **Batch Size:** 32
- **Epochs:** 50 (early stopping patience=10)
- **Weight Decay (L2):** 0.0005
- **Augmentation:** Mosaic (80%), Mixup (10%), rotation (±15°), brightness/contrast (±20%)

---

3. Data & Training

Training Data Characteristics

- **Source:** Custom satellite imagery of Indian rooftop solar installations
- **Period:** 2024-2025
- **Geographic Coverage:** Pan-India (28 states)
- **Satellite Sources:** Sentinel-2, Planet Labs, commercial providers
- **Total Samples:** ~5,000 image pairs
- **Cloud Filter:** $<$ 20% cloud cover
- **Resolution:** 10m/pixel (standard), 1-2m/pixel (high-resolution)

Data Preprocessing

- Image normalization (pixel values $\rightarrow$ [0, 1])
- Geo-registration to standard coordinates
- Cloud masking (>20% filtered)
- Geographic stratification (70/15/15 split)
- Class balancing via augmentation
- 15% metadata interpolation
- 3% corrupted images removed

Annotations

- **Method:** Manual with expert verification
- **Format:** YOLO (normalized coordinates)
- **Classes:** Solar panels, partial visibility, shadows, related structures
- **Inter-annotator Agreement:** 0.89 (Cohen's Kappa)
- **Ground Truth Validation:** Field surveys (n=500)

---

4. Performance Analysis

Geographic Performance Variation

Region	mAP@0.5	Precision	Recall	Status
Northern India	0.847	0.782	0.801	
Western India	0.832	0.768	0.795	
Southern India	0.816	0.751	0.775	
Eastern India	0.798	0.721	0.743	
Northeast India	**0.761**	**0.684**	**0.712**	

Disparity: 8.6% gap between best and worst performing regions (fairness concern)

Installation Type Performance

Type	Precision	Recall	F1-Score
Large commercial (>100 kW)	0.89	0.91	0.90
Medium residential (5-25 kW)	0.78	0.81	0.79
Small residential (<5 kW)	0.68	0.70	0.69

Scale Bias: 21% performance gap between large and small installations

Common Failure Modes

1. **Shadow Confusion (12%)** – High sun angles cause false detection
2. **Partial Visibility (18%)** – Obstructed panels (>20% covered)
3. **Similar Structures (15%)** – Confusion with water tanks, AC units
4. **Small/Distant Panels (20%)** – Sub-pixel detection limits
5. **Weathered Panels (10%)** – Aged/soiled panels (accuracy drops to 0.52)

---

5. Known Limitations

Model Limitations

- **Satellite Dependency:** Performance tied to image quality and revisit frequency (5-10 days)
- **Cloud Cover:** >20% cloud cover causes failures
- **Temporal Scope:** Trained on 2024-2025 data only
- **Scale Limits:** Best for 5-100 kW installations; poor for <1 kW or >500 kW
- **Geographic Scope:** India-specific; transfer to other countries not validated
- **Explainability:** YOLOv11 black-box nature limits interpretability

Critical Failure Cases

- **High Reflectance Roofs:** 15-20% false positive rate
- **Dense Clusters:** 8-12% undercount in >200-panel facilities
- **Wet/Soiled Panels:** 22% accuracy drop post-rainfall (24-48hr)
- **Extreme Shadows:** 35% failure rate when sun angle <30°
- **Obstructed Views:** 45% miss rate for >20% obstructed panels

---

6. Fairness & Ethical Considerations

Identified Risks

Risk	Severity	Mitigation
Geographic discrimination (NE India)	**HIGH**	Regional oversampling, human review for <0.70 confidence
Scale discrimination (<5kW households)	**HIGH**	Sub-model development, confidence-based routing
False rejections (8.2% false positive)	**MEDIUM**	Appeal mechanism, manual review mandatory
Seasonal disparities (monsoon/winter)	**MEDIUM**	Temporal augmentation, seasonal analysis
Privacy concerns (satellite imagery)	**MEDIUM**	Data anonymization, secure storage

Fairness Metrics

- **Demographic Parity:** 6.3% difference (acceptable)
- **Equalized Odds:** 7.1% difference (passes threshold)

- **Predictive Parity:** 4.8% difference (passes threshold)
- **Disparate Impact Ratio:** 0.76 (slightly below 0.80 threshold)

Recommendations

- **Do NOT** use as sole decision-maker; implement human oversight for all rejections
- **DO** implement confidence-based routing (auto-approve only for score 0.85)
- **DO** conduct quarterly fairness audits across regions and scales
- **DO** maintain 100% manual review for auto-rejections
- **DO** establish appeals process with clear documentation

---

7. Intended Use & Limitations

Primary Use Cases

- Automated verification of solar installations in MNRE subsidy program
- Fraud detection (fake installations, spoofed locations, tampering)
- Batch processing of satellite imagery datasets
- Decision support for subsidy officers

NOT Designed For

- Sole decision-maker in high-stakes approvals
- Non-solar aerial imagery without retraining
- Real-time streaming without buffering
- Applications outside Indian solar context
- Installations <5 kW or with $>20\%$ obstruction

User Groups

- Government subsidy verification officers
- Solar installation companies
- Administrative authorities
- Data scientists maintaining the system

---

8. Maintenance & Monitoring

Retraining Schedule

- **Quarterly:** Full retraining with 1,000+ new samples
- **Bi-annually:** Fairness audit and regional analysis
- **Annually:** Architecture upgrade and hyperparameter tuning

Retraining Triggers

- mAP@0.5 drops below 0.81 (3% degradation)
- Regional performance gap exceeds 10%
- Accuracy drop >5% on out-of-distribution imagery
- New panel technologies not detected
- Satellite imagery provider changes

Monitoring Metrics

- **Real-time:** Latency (<100ms), throughput (>15 img/sec), confidence distribution
- **Weekly:** Accuracy, error rates, system health
- **Monthly:** Regional performance breakdown, drift detection
- **Quarterly:** Comprehensive fairness audit, bias analysis

Model Deprecation

- **Trigger:** mAP@0.5 < 0.80 for 2 consecutive quarters
- **Transition:** 6-month dual model operation
- **Replacement:** Version 2.0 (multi-temporal + regional sub-models)
- **Support:** Critical patches until Dec 2026

---

9. Key Statistics & Benchmarks

Training Summary

- **Total Training Time:** 2.8 hours (50 epochs)
- **Hardware:** NVIDIA RTX 4090 (24GB VRAM), Intel i9-13900K
- **Convergence:** Epoch 35 (loss stabilized after this)
- **Overfitting:** Minimal (validation loss gap 0.13)
- **Cross-Validation:** Std dev = 0.031 (5-fold) – robust performance

Test Set Breakdown

- **True Positives:** 5,891 panels correctly detected
- **False Positives:** 1,942 incorrect detections
- **False Negatives:** 1,987 missed panels
- **Precision-Recall Curve:** AUC = 0.824 @ IoU=0.5

Inference Performance

Device	Latency	Throughput
NVIDIA RTX 4090	45ms	22 img/sec
Intel i9-13900K CPU	180ms	5.5 img/sec

Device	Latency	Throughput
NVIDIA Jetson Orin	85ms	11.8 img/sec
AWS SageMaker	52ms	19 img/sec

---

10. Documentation & Resources

Files in Deliverables

- **Model Weights:** `trained_model_file/yolov11_solar_panel_detector.pt` (5MB)
- **Pipeline Code:** `pipeline_code/main_pipeline.py` (batch inference runner)
- **Training Logs:** `model_training_logs/training_metrics.csv` (50 epochs)
- **Sample Predictions:** `prediction_files/verification_predictions_*.json`
- **Environment:** `environment_details/requirements.txt`, `environment.yml`
- **Visualizations:** `artefacts/` (detection samples, training curves, metrics)

Running the Pipeline

```
# Quick start
conda env create -f environment_details/environment.yml
conda activate surya-saathi
python pipeline_code/main_pipeline.py
```

API Integration

```
POST /api/v1/applications/submit
{
  "latitude": 28.6139,
  "longitude": 77.2090,
  "declared_panel_count": 20,
  "submission_date": "2025-12-12"
}
# Returns: confidence_score, panel_count, qc_status
```

---

11. References & Citation

Primary References: - Ultralytics YOLOv11: "YOLOv11: Faster, Better, Stronger" (2024) - Solar Panel Detection: Chen et al., "Deep Learning-Based Rooftop Solar Panel Detection" (IEEE 2023) - Fairness in ML: Buolamwini &

Gebru, "Gender Shades" (FAT* 2018) - India's Solar Program: MNRE Official Documentation (2024)

Citation:

```
@model{suryasaathi2025,
  title={YOLOv11-based Solar Panel Detection for Fraud Prevention},
  author={AlgoMinds Team},
  year={2025},
  note={Surya Saathi: AI-Powered Solar Subsidy Verification System}
}
```

Model Card Metadata

Field	Value
Last Updated	December 12, 2025
Next Review	March 12, 2026
Current Status	Active in Production
Maintenance	Actively maintained by AlgoMinds Team
Contact	ayushnp@gmail.com
Repository	github.com/ayushnp/AlgoMinds-SuryaSaathi

Surya Saathi: Preventing fraud and accelerating India's green energy transition using AI and satellite intelligence.

For comprehensive documentation, see the full Model Card PDF and README.md in the deliverables folder.