Remote Sensing Assessment and Monitoring of Distributed Rooftop Solar Panel Arrays

A report on methods for Computer Vision extraction and Photovoltaic energy forecasting

Alejandro S. Vega-Nogales, Data Scientist @ Maxar, CCOM MS Student
SIN: 801-13-7956 github: avega17 email: alejandro.vega1@upr.edu
UPRRP - Computer Science Department
CCOM 6050: Final Project – Midterm Project Report

Contents

1 Introduction

1.1 General Objective

Regional-scale surveys of distributed *rooftop PV solar panels*, and **site-level, short-term forecasting** of *solar irradiance* used to estimate local PV power generation.

This assessment will primarily produce counts of installed systems, estimates of surface area, estimated PV generation potential while monitoring will provide forecasts of local solar irradiance which are used to model estimates of PV energy generation at specific sites.

The work developed for this thesis proposal is particularly inspired by and seeks to build upon the work realized, methodologies outlined, research gaps identified, and future work outlined in[?],[?],[?],[?],[?],[?],[?], and[?].

1.2 Specific Objectives

1. Survey of rooftop Solar Panel *Arrays* via Small Object Detection and Instance Segmentation using very-high-spatial-resolution imagery ($\leq$ 1 meter/pixel) of installation locations sourced from a **global** inventory of labeled PV arrays collated from multiple scientific dataset publications [add citations]
2. *Near-term site-level* **forecasting of solar Global Horizontal Irradiance** used to estimate PV power generation via data fusion with Spatio-temporal context from:
   - Time-series frames from geostationary satellite sensors with coarse spatial resolution, but high-temporal resolution (intra-hour scans)
   - Historical Solar irradiance (e.g. NREL's National Solar Radiation Database)
   - Any available PV power generation ground-truth time series

1.3 Contribution Goals

1. Measure performance impact of modern Computer Vision architectures (e.g. Vision Transformers, Swin Transformer, CNN-Transformer hybrids) using much higher resolution imagery than most published research
2. Measure impact of use of PV-specific spectral indices and specifically the benefits of including NIR + SWIR bands
3. Develop a solution that consciously tackles the challenges identified in[?] for evaluating and performing comparisons of different remote sensing solar array assessment methodologies (distribution drift, test data quality, level of spatial-aggregation, and proprietary data)

---

2 Background

2.1 Preliminaries

2.1.1 Remote Sensing and Geospatial Data

Remote sensing (RS) refers to the process of acquiring images and data of the planet's surface using a variety of remote sensors in satellite and aerial vehicles (add citations from my prev paper). These sensors analyze electromagnetic radiation reflected or emitted from objects on the Earth's surface, which is then processed to extract information about the objects and their properties. Besides traditional electro-optical (ie panchromatic and 3-channel RGB images (cite)) modern RS employs a variety of sensors and modalities such as multispectral (four or more non-overlapping bands in the electromagnetic spectrum), hyperspectral (more than 100 narrow bands), and even active sensors such as microwave altimeters or Synthetic Aperture Radar (SAR) that emit their own radiation and measure the reflected signal to "see" at night and through atmospheric obstructions like clouds and fog (cite myself or refs from prev paper). Analysis of source data from these remote sensors must handle a variety of nuanced characteristics of each sensor and the produced data. For our purposes and the scope of this report, we will limit our discussion to the following characteristics:

- Spatial resolution: The level of detail in an image, typically measured in meters per pixel.
- Spectral resolution: The ability of a sensor to distinguish between different wavelengths of light.
- Temporal resolution: The frequency at which a sensor captures data over the same location.
- Radiometric resolution: The sensitivity of a sensor to detect variations in intensity, often represented by bit depth.
- Earth Observation (EO) data types:
  1. Raster data: Gridded data representing continuous surfaces.
  2. Vector data: Discrete data represented as points, lines, or polygons.
  3. Time series data: Sequential data capturing changes over time.
  4. Geospatial data cubes: Multidimensional arrays combining spatial, temporal, and spectral dimensions.

A fundamental challenge in RS is the tradeoff that an individual sensor's orbit and design must make between spatial, spectral, and temporal resolution.

2.1.2 Computer Vision

Computer Vision (CV) is a subfield of Artificial Intelligence and a discipline that deals with the problem of interpreting and extracting meaningful information from images (cite myself or refs from prev paper) in a manner similar to human vision. This field has seen significant advancements in recent years, particularly with the rise of deep learning methods and the spread of large, labeled datasets for many applications. RS and EO has seen an explosion of interest, publications, and datasets for DL methods since 2015 (cite EO paper from cloud computing course). Relevant CV tasks for our purposes include:

1. Object Detection: Identifying and localizing specific object classes (e.g., PV panels) within an image with (georeferenced, in our case) bounding boxes.
2. Semantic Segmentation: Classifying each pixel in an image into predefined categories (e.g., PV panel array, rooftop, vegetation, background).
3. Instance Segmentation: Similar to semantic segmentation, but differentiates

between *individual* instances (e.g., distinguishing between different PV panel arrays).

Some relevant CV architectures include Convolutional Neural Networks (CNNs), Vision Transformers (ViTs), Generative Adversarial Networks (GANs), and CNN-Transformer hybrids. (detail in later draft...)

2.1.3 Data Fusion

1. Definition: combining data from multiple sources to achieve improved information quality or inference compared to using sources individually[?]
2. (Geospatial) Coherence[?]
3. Remote Sensing Fusion types[?][?]
   - Spatio-Spectral Fusion: Enhancing spectral resolution using spatial information (e.g., Pansharpening) or vice versa.[?][?]
   - Spatio-Temporal Fusion: Combining high spatial resolution/low temporal frequency data with low spatial resolution/high temporal frequency data to generate a fused dataset with high resolution in both domains.[?] **This is the core fusion type for the second subproblem of this project**.
4. Data Fusion Abstraction Levels[?]
   - Early fusion (pixel/signal level): Combining raw sensor data directly
   - Intermediate fusion (feature level): Extracting relevant features from each sensor first and then combining them
   - Late fusion (symbol/decision level): Each source produces an independent decision, which is then combined to produce a final output
5. A review from Castanedo[?] also categorizes fusion based on the relationship between data sources (complementary, redundant, cooperative)
6. The rise of Deep Learning (DL) has significantly impacted the field, with reviews listing CNNs, LSTMs, GANs, and Transformers as the most relevant architectures for data fusion[?][?].

2.2 Problem Statement and Motivation

The global response to climate change requires our civilization to implement a rapid transition to renewable energy sources in an effort to decarbonize the energy sector as much as possible. The fastest growing renewable energy source is by far *photovoltaic* (PV) solar energy. Growing at over 41% per year since 2009[?], PV solar energy has far outpaced other renewable energy sources such as concentrated solar power (CSP), hydroelectric storage, geothermal energy, and, to a lesser extent, wind energy. This rapid growth has led to significant progress in international development goals such as the United Nations Sustainable Development Goals (SDG) which, among other things, addresses the population level needs for clean and affordable energy, and actions to tackle climate change and its impacts[?].

Historically, one of the most effective means to ensure global cooperation and uniform implementation of environmental and energy policies has been through the establishment of multi-lateral international agreements such as the Montreal Protocol, the Kyoto Protocol, and, most recently, the Paris Climate Accords signed by (add figure of % of countries signed and ref) countries. This agreement resulted in *non-binding* commitments to reduce greenhouse gas emissions and goals to limit global warming to 1.5 - 2.0 degrees Celsius above pre-industrial levels. Notably, 87% of these Nationally Determined Contributions (NDCs) aim to increase the share of renewable energy in their energy mix with half specifying empirical generation figures[?]. To maintain any semblance of international accountability to these commitments, these NDCs require extensive and accurate monitoring and validation of any expansion of renewable energy infrastructure.

To address this *international policy problem* there has been a variety of efforts to assess the progress of each renewable energy source leading to recurring reports, inventories, assessments, and databases at the global, regional, and national levels. Many of these efforts prove insufficient due to only providing aggregated summaries or statistical extrapolations at the national level (IRENA, IEA, BP), or being limited to specific regions (e.g. Europe, North America, China). Particularly in the case of PV solar energy, there has been very significant progress over the last decade in the development of remote sensing and deep learning methods to assess the distribution of PV solar energy at the site, regional, and global levels (cite the many PV assessment papers). Most recently, a collaboration between Microsoft, the Nature Conservancy, and Planet Labs has produced *"Global Renewables Watch"*, a public dataset of industrial, commercial, and utility-scale PV solar energy and wind energy installations showcasing the potential of using high-resolution satellite imagery and deep learning methods to assess the distribution of renewable energy infrastructure at the planetary scale.

However, for some countries, land comes at a premium and concentrated large-scale solar energy installations are either not economically or politically feasible due to not being the most efficient use of land and other land use conflicts. This leads to countries where distributed, small-scale rooftop or building-integrated PV solar energy installations make up a significant portion of the total installed capacity which is not captured in assessment limited to large-scale installations ($\geq$ 10KW or area-figure-threshold)[?][?]. Additionally, in the case of intermittent renewable energy sources, there is a second *technical problem* that also calls for assessments and inventories at a smaller spatial **and** temporal scale.

The intermittency of renewable energy sources such as wind and solar energy means that energy grid operators require accurate short-term forecasts of energy generation to ensure that supply and demand are balanced. PV solar energy is particularly dependent on local weather conditions and received solar irradiance which can both be monitored using remote sensing methods and local weather stations. There has been recent work in the use of geostationary satellites with high temporal resolution to provide short-term forecasts of solar irradiance at the site level[?], the use of spectral reflectance features in Computer Vision (CV) models for PV solar array detection[?], and advances in spatio-temporal data fusion methods to improve the temporal resolution of very-high-resolution (VHR) satellite imagery[?].

This project looks to outline the work required to address these interconnected challenges by tackling two core subproblems leveraging remote sensing data and advances in deep learning methods:

2.2.1 [CCOM6120] Subproblem 1:

Computer Vision detection and assessment of distributed rooftop PV Systems

The first fundamental challenge lies in the automated, accurate, and scalable identification and geometric characterization of distributed rooftop PV panels using very-high-resolution (VHR) multispectral satellite imagery. Generating reliable, up-to-date inventories of these small, dispersed assets is crucial for granular PV potential assessments[?][?], infrastructure planning, monitoring deployment rates against policy goals[?], and producing the georeferenced geometry data required for accurate site-level solar irradiance forecasting and PV energy generation estimates[?].

2.2.2 [CCOM6050] Subproblem 2:

Remote Sensing data fusion for Solar Irradiance time series forecasting

The second major challenge is generating ~~high-spatiotemporal-resolution surface reflectance~~ time series of estimated solar irradiance at lower spatial resolutions through the fusion of multispectral satellite data from sensors with complementary remote sensing characteristics:

- **Fine spatial** ($\leq 10m$) but **coarse temporal** (days or weeks) resolution:
  While VHR imagery provides spatial detail, its infrequent revisit limits its use for monitoring dynamic processes like cloud cover, which governs solar irradiance variability (cite and rephrase).
- **Coarse spatial** ($\geq 100m$) but **fine temporal** (mins. or hr.) resolution:
  Conversely, geostationary (e.g., GOES-R) or wide-swath polar-orbiting satellites (e.g., MODIS, Sentinel-3) offer high temporal frequency but lack the spatial resolution needed for site-specific analysis of distributed PV.

3 Methodology

Here you will describe your:

- Experiments done and pending
- Tools, software, code
- Data available (Description, source)
- Results (if any)

Create a subsection for each part described in the methodology. These subsections can be modified based on the work or progress made during the current month. Additionally, you may create separate sections if needed, for example, to distinguish results based on different strategies. This template is flexible and can be adapted for better presentation.

3.1 Datasets and Processing

3.1.1 RS imagery and locations of PV arrays

- Dataset 1 - Description, source
- Dataset 2 - Description, source

3.1.2 RS Datasets for STF

- Dataset 1 - Description, source
- Dataset 2 - Description, source

3.1.3 Preprocessing steps

3.2 Implementation Details

3.2.1 Python Libraries and other software

- *Python 3.10* - programming language
- *PyTorch 2.0* - deep learning framework
- *jupyterlab* - web-based interactive development environment for Jupyter notebooks

- *numpy + xarray* - array processing
- *Geopandas* - Geospatial data manipulation and analysis
- *duckdb* - an in-process SQL OLAP database management system
- *cubo* - a library for working with cuboids (3D arrays) in Python
- *openeo* - python client for the OpenEO API
- *Rasterio* - Raster data reading and writing
- *GDAL* - Geospatial data abstraction library
- *mamba* - a faster alternative to the conda package manager
- *rastervision* - a framework for computer vision and deep learning in remote sensing
- *PyTorch Lightning* - a lightweight wrapper for PyTorch to help with training and testing
- *TorchGeo* - a library for deep learning on geospatial data which includes datasets and pretrained models
- *super-gradients* - a library for transfer learning and training of SOTA CV models
- *data-gradients* - a library designed for computer vision dataset analysis and visualization
- *torchmetrics* - a library for computing metrics in PyTorch

3.2.2 Hardware and Compute Resources

3.2.3 Code Repository

At the end of the semester, the private Github repo where this report and associated code is hosted will be made public.

3.3 Evaluation

- Goal(s): Assess accuracy of reconstructed time series
- Metrics (including their normalized versions):
  - RMSE (Root Mean Square Error)
  - MAE (Mean Absolute Error)
  - MAPE (Mean Absolute Percentage Error)
  - MBD (Mean Bias Deviation)
  - R2 (Coefficient of Determination)
  - PSNR (Peak Signal-to-Noise Ratio)
- Baseline(s):
  - Persistence model
  - Clear sky components

3.4 Results

4 [CCOM6050] Data Fusion Algorithms

Here we will go into more detail and implement one traditional data fusion algorithm and another using Deep Learning methodology.

4.1 Algorithm Design

4.1.1 Statistical STF algorithm candidates

(add citations)

5 Literature Review

5.1 [CCOM6050] Data Fusion Reviews

5.1.1 2019-RS-DF-Multisource

- Title: "Multisource and Multitemporal Data Fusion in Remote Sensing: A Comprehensive Review of the State of the Art"
- Authors: Pedram Ghamisi; Behnood Rasti; Naoto Yokoya; Qunming Wang; Bernhard Hofle; Lorenzo Bruzzone
- Year: 2019
- Link/Source: 10.1109/MGRS.2018.2890023
- Main objective:
- Methodology used:
- What I got out of the paper:
  - "bird's eye view" of contributions in:
    - pansharpening and resolution enhancement
    - multitemporal data fusion
  - Traditionally in Remote Sensing (RS), there are four dimensions that provide information:
    - Spatial: 2D (x,y) coordinates
    - Temporal : 1D (t) time
    - Spectral: 1D (λ) wavelength
    - Radiometric: 1D (r) numerical precision in a pixel's radiance/reflectance/intensity values
  - RS measurement process is explored via describing the four components of a "physical model":
    - Scene model: defines the subject of interest
    - Atmosphere model: defines the transform of Electromagnetic Spectrum (EMS) from surface to sensor
    - Sensor model: defines the measurement process (e.g. signal-to-noise ratio (SNR), sweep time, nadir angle, number of bands, etc.)
    - Image model: defines the sampling process (e.g. pixel size, spatial resolution, etc.)
  - **"All data fusion methods attempt to overcome the above measurement and sampling processes"**
    "Understanding these [differences in] measurement and sampling processes is... key

to characterizing methods of data fusion" $\implies$ can be approached as a reconstruction problem...

- Spatio-Spectral data fusion
  * Pansharpening: a specific case of spatio-spectral fusion where a high-resolution panchromatic image is fused with a lower-resolution multispectral image.
    This fusion process attempts to preserve the spectral information of the multispectral image while enhancing the spatial resolution for *all* bands.
  * Most methods can be categorized into:
    1. Component substitution: spectrally transforms MS data into a new feature space to separate spatial and spectral information,
       then substitutes the spatial component is substituted for the high-resolution panchromatic image while using histogram matching to adjust the PAN intensity to match the original spectral information.
       · Pros: *high spatial fidelity*; *low computational cost*; *robust against misregistration errors (i.e. spatial misalignment)*
       · Cons: *suffer from global spectral distortion*; *lower spectral fidelity*
       · Methods include:
       (a) **Intensity-Hue-Saturation (IHS)**
       (b) Principal Component Analysis (PCA)
       (c) Gram-Schmidt (GS)
    2. Multiresolution analysis: extracts spatial high-frequency components from the panchromatic image and injects with coefficients into the low-resolution multispectral image.
       · Pros: Spectral consistency (e.g. no spectral loss)
       · Cons: *Higher computational complexity and cost*
       · Methods include:
       (a) Box filtering
       (b) Gaussian filtering
       (c) Bilateral filtering
       (d) Wavelet transform
       (e) Curvelet transform
    3. Geostatistical analysis: can preserve the spectral properties of the original coarse image while enhancing the spatial resolution.
       · Pros:
       · Cons:

  - Methods include: several types of *kriging* (spatial interpolation) which are used to estimate unmeasured values based on neighboring values

4. Subspace representation: uses a subspace spanned by a set of basis vectors to analyze intrinsic spectral characteristics
    - Pros:
    - Cons: *High computational cost*; *can introduce spectral artifacts*
    - Methods include: Bayesian analysis, Matrix Factorization, Spectral Unmixing
5. Sparse representation: captures the spectral signatures of materials in an image patch using a sparse representation with a few basis vectors
    - Pros: *Efficient storage and processing*; *high spectral fidelity*; *sparse representation can be used in downstream CV tasks*
    - Cons: *Targeted for HSI-MSI fusion*; *very high computational cost*
    - Methods include: Hierarchical Pyramid models; extraction of spectral signatures;

* Some quantitative evaluation metrics:
  1. Peak Signar-to-Noise Ratio (PSNR): evaluates the quality of a reconstructed image by comparing the maximum possible power of a signal to the power of corrupting noise
  2. Spectral Angle Mapper (SAM): determines the similarity between a transformed image spectrum and a reference spectrum by calculating the angle between their vector representations in the spectral space
  3. ERGAS (*erreur relative globale adimensionnelle de synthèse*): measures image quality in terms of the per-band normalized mean error between the fused image and the reference image
  4. Q^{2n}: a global reconstruction quality index

- Spatio-temporal data fusion
  * "a technique to blend fine spatial resolution, but coarse temporal resolution data with fine temporal resolution, but coarse spatial resolution data"
  * **Most methods are based on the strong assumption of no abrupt changes in Land Cover and Land Use (LCLU) across time** (ok for our work if we focus on rooftops)
  * Methods mentioned:
    1. STARFM (Spatial and Temporal Adaptive Reflectance Fusion Model)
    2. ESTARFM (Enhanced STARFM)
  * Common sensor pairings:
    1. MODIS (daily; 0.25km-1km) + Landsat (16 days; 30m)

2. MODIS + Sentinel-2 (5 days; 10m)
3. GOES-R (5-60 min; 0.25-2km) + ??

- Key findings or contributions for our topic:
  - Outlines the centrality of Coherence for Data Fusion with RS data:

    "Data fusion is made possible because each dataset to be fused represents a different view of the same real world defined in space and time (generalized by the scene model), with each view having its own measurable properties, measurement processes, and sampling processes."

    $\Longrightarrow$ "one should expect some level of **coherence** between the real world (the source) and the multiple datasets (the observations)"
  - **Spatio-temporal fusion** was identified as the relevant type of fusion for our work:

    "A large focus of attention currently is on the specific problem that arises from the trade-off in remote sensing between spatial resolution and temporal frequency; in particular the fusion of coarse-spatial/fine-temporal-resolution with fine-spatial/coarse-temporal-resolution space-time datasets"
  - For an individual remote sensing platform, "there always exists a trade-off between spatial resolution and temporal resolution (revisit time)". This can be worked around with a constellation of multiple sensors.
  - A remaining major issue is how to conduct fair comparisons of performance and accuracy of different methods.

5.1.2 2024-DL-Data-Fusion

- Title: "A Comprehensive Review on Deep Learning-Based Data Fusion"'
- Authors: Mazhar Hussain; Mattias O' Nils; Jan Lundgren; Seyed Jalaleddin Mousavirad
- Year: 2024
- Link/Source: 10.1109/ACCESS.2024.3508271
- Main objective:
- Methodology used:
- Key findings or contributions for our topic:

5.2 [CCOM6050] Data Fusion Datasets

- Title:
- Authors:
- Year:
- Link/Source:

- Main objective of the paper:
- Methodology used:
- Number of annotations:
- Date range for annotations:
- Annotation locations:
- Key findings or contributions:

5.3 PV imagery and locations Datasets

5.3.1 2023-SDG-Maxar-PV-dataset

- Title: "A solar panel dataset of very high resolution satellite imagery to support the Sustainable Development Goals"
- Authors: Cecilia Clark [ex-Maxar]; Fabio Pacifici[Maxar]
- Year: 2023
- Link/Source: 10.1038/s41597-023-02539-8 — imagery — annotations
- Main objective of the paper: To provide a VHR satellite imagery dataset of annotated, primarily residential, solar panels to support UN's Sustainable Development Goals, and further improve solar panel detection models.
- Methodology used:
  Obtained 31 cm resolution satellite imagery and applied proprietary HD processing to generate 15.5 cm resolution imagery used for panel detection. Labels created with Object Detection model (YOLT) and validated by human annotators.
- Number of annotations: 2,542
- Date range for annotations: 2020/09/18
- Annotation locations: Southern Germany
- Key findings or contributions:
  The dataset is designed to support small object detection and focuses on annotated, primarily residential, solar panels. Includes paired native resolution (31 cm) and HD (15.5 cm) satellite imagery.

5.3.2 2021-global-PV-inventory

- Title: "A global inventory of photovoltaic solar energy generating units"
- Authors: L. Kruitwagen; K. T. Story; J. Friedrich; L. Byers; S. Skillman; C. Hepburn
- Year: 2021
- Link/Source: 10.1038/s41586-021-03957-7 — PV Labels — Code Repo
- Main objective of the paper: To provide a global inventory of commercial, industrial, and utility-scale PV installations ($\geq 10kW$ nominal generation capacity)

- Methodology used: Machine Learning pipeline (series of CNN + RNN models), OpenStreetMap annotations, heuristic filters, negative sampling (non solar panel objects)
- Number of annotations: 68,661 facilities [36,882 (Sentinel-2), 38,541 (SPOT)]
- Date range for annotations: 2016/06/01 - 2018/09/30
- Annotation locations: Global (131 countries)
- Key findings or contributions: The dataset expands previous publicly available data by $> 4x$. Provides an estimate of global installed generating capacity: $423(\pm 75) GW$. Mentions importance of spectral signature of PV panels and extraction of spectral features. Provides implementation repo. "The pipeline has two stages: an initial global search designed to maximize installation recall, followed by a process to remove false positives and estimate installation dates."

5.3.3 2020-UK-Solar-PV

- Title: "A harmonised, high-coverage, open dataset of solar photovoltaic installations in the UK"
- Authors: Dan Stowell; Jack Kelly; Damien Tanner; Jamie Taylor; Ethan Jones; James Geddes; Ed Chalstrey
- Year: 2020
- Link/Source: 10.1038/s41597-020-00739-0 — PV Labels — Code Repo
- Main objective of the paper: To create an open geographic data source for solar PV, suitable for intra-nation (UK in this case) analysis using machine vision and PV forecasting.
- Methodology used: OpenStreetMap, Crowdsourcing, deduplication via spatial clustering
- Number of annotations: over 260,000 (over 255K separate installations; 1067 large solar farms)
- Date range for annotations: Data was collected up to September 2020
- Annotation locations: United Kingdom
- Key findings or contributions: It includes a large number of small-scale domestic installations, which were typically poorly documented. Provides detailed metadata and location geometries. Discusses challenges of data reconciliation and deduplication which is relevant to our work that will merge multiple datasets. Includes a GUI for data visualization and validation which will likely be relevant or useful for our work

5.3.4 2023-US-large-PV-EIA

- Title: "Georectified polygon database of ground-mounted large-scale solar photovoltaic sites in the United States"
- Authors: K. Sydny Fujita; Zachary H. Ancona; Louisa A. Kramer; Mary Straka; Tandie E. Gautreau; Dana Robson; Chris Garrity; Ben Hoen; Jay E. Diffendorfer
- Year: 2023