Deep Learning-Based Monitoring of Solar Farm Dynamics in Ba Ria–Vung Tau Using Multi-Temporal Sentinel-2 Imagery

Nguyen Van Thang[1], Le Hong Anh[1], Nguyen Thi Mai Dung[1], and Tran Mai Huong[1]

Faculty of Information Technology,
Hanoi University of Mining and Geology, Vietnam
{nguyenthimaidung}@humg.edu.vn

Abstract. This study presents a remote sensing framework for annual solar farm mapping and spatiotemporal change monitoring in Ba Ria–Vung Tau, Vietnam. Using Sentinel-2 multi-temporal imagery (2020–2025), we generated cloud-reduced seasonal composites based on seven spectral bands ($B2, B3, B4, B6, B8, B11, B12$) and a photovoltaic spectral index (PVSI). Data from 2020–2022 were used for model training, while 2023–2025 composites were used for province-wide inference and change analysis. Three deep learning architectures, U-Net, LinkNet, and DeepLabV3+, were evaluated, with LinkNet achieving the best overall performance (IoU = 90.74%, F1-score = 93.32%). An ablation experiment further showed that PVSI provided modest but consistent gains in segmentation quality. A post-segmentation pipeline and interannual differencing were then employed to quantify newly emerged, persistent, and expanded solar farm areas. Results demonstrate that medium-resolution Sentinel-2 imagery combined with deep semantic segmentation provides a reproducible and practically useful basis for annual monitoring.

Keywords: Solar farm monitoring, Sentinel-2, Semantic segmentation, Ba Ria–Vung Tau, Deep learning

1 Introduction

Vietnam has undergone a significant energy transition over the past decade, with solar photovoltaic (PV) capacity expanding from negligible levels to several gigawatts within a short policy-driven window [1, 2]. This rapid deployment of utility-scale and rooftop solar systems has been particularly concentrated in southern coastal provinces, where irradiance conditions are favorable and industrial land is available. Ba Ria–Vung Tau (BRVT) province, located in the southeastern coastal region of Vietnam, has emerged as a notable zone for solar farm development, alongside its established roles in petroleum extraction, port logistics, and industrial manufacturing.

Despite the scale and pace of this transformation, systematic spatial monitoring of solar farm distribution and dynamics in BRVT remains limited. Decision-makers and planners typically rely on administrative records or periodic field

surveys, which lack the spatial explicitness and temporal granularity required for informed land-use management. Remote sensing offers a scalable and consistent alternative: satellite imagery provides repeatable, georeferenced observations that can document infrastructure changes at annual or sub-annual intervals without field access constraints.

Sentinel-2, operated by the European Space Agency as part of the Copernicus programme, provides multispectral imagery at 10–20 m spatial resolution with a revisit period of approximately five days. This temporal density and spatial resolution make it well-suited for monitoring medium-to-large solar installations at the provincial scale. However, converting raw time-series imagery into meaningful change maps requires a structured pipeline addressing cloud contamination, inter-annual radiometric variation, and the semantic complexity of distinguishing solar farms from spectrally similar surfaces.

Deep learning-based semantic segmentation has become an effective approach for extracting land cover classes from remote sensing imagery [3, 4]. Encoder–decoder architectures such as U-Net [5], LinkNet, and related models have shown strong performance in infrastructure and land cover mapping. However, most existing studies focus primarily on single-date detection, while comparatively less attention has been given to year-by-year monitoring of solar farm dynamics, including persistence, expansion, and newly emerged installations. In addition, the effectiveness of medium-resolution Sentinel-2 inputs enhanced with photovoltaic-specific spectral features remains insufficiently explored.

To address this gap, this paper presents a monitoring-oriented framework that combines multi-temporal Sentinel-2 imagery, photovoltaic-oriented feature construction, deep semantic segmentation, annual binary mapping, and interannual change analysis for Ba Ria–Vung Tau province. The framework uses seven Sentinel-2 spectral bands together with a photovoltaic spectral index, and compares segmentation architectures prior to province-wide deployment for monitoring from 2023 to 2025.

The main contributions of this paper are as follows:

1. A practical and reproducible workflow for annual solar farm mapping in Ba Ria–Vung Tau using Sentinel-2 seasonal composites, multispectral inputs, and deep semantic segmentation.
2. A comparative evaluation of three segmentation architectures U-Net, LinkNet, and DeepLabV3+ together with an ablation study assessing the contribution of the PVSI channel for medium-resolution solar farm mapping.
3. A multi-year monitoring framework (2023–2025) that quantifies the spatial evolution of solar farm infrastructure at the provincial scale through annual footprint generation and interannual change analysis.

2 Related Work

2.1 Remote sensing of Solar PV infrastructure

Remote sensing for solar PV infrastructure has evolved rapidly with the increasing availability of high-resolution imagery and advances in deep learning

based semantic segmentation. Early work relied on thresholded spectral indices, object-based image analysis, or handcrafted texture features applied to very high-resolution aerial or satellite imagery, which achieved good local performance but required substantial manual parameter tuning and were difficult to scale to large regions. More recent efforts have shifted toward training convolutional neural networks (CNNs) on large image archives to detect and delineate solar PV arrays automatically, enabling regional to global mapping at regular intervals [6–8].

Semantic segmentation architectures such as U-Net and DeepLabV3+ have become the dominant approach for PV plant detection from medium and high resolution optical data. CNN-based models have been used to generate global or continental-scale maps of utility-scale solar farms, often using PlanetScope or Sentinel-2 imagery as input and then vectorizing the predicted masks into geospatial polygons [7, 8]. Several studies report that encoder–decoder CNNs remain highly competitive, and in many cases superior, to contemporary Vision Transformers when trained on medium-resolution multispectral data, particularly under realistic constraints on training data volume and compute resources [9]. These findings are consistent with broader remote sensing results indicating that CNNs provide robust performance and stable optimization for large-area land-cover mapping from Sentinel-2 [10]. In parallel, recent work has begun to explore time-series-aware segmentation models and temporal feature engineering to better capture the phenological and illumination variability surrounding solar farms [11].

Transformer and hybrid architectures, including SegFormer and DeepLab variants with attention modules, have demonstrated strong performance in complex urban and heterogeneous landscapes [11]. However, their advantages tend to be most pronounced when very large, diverse training sets are available and when high computational budgets can be devoted to training and inference. For operational mapping of solar farms at provincial to national scale using medium-resolution multispectral imagery, convolutional encoder–decoder models therefore remain a practical choice, offering a favorable balance between accuracy, model complexity, and deployment cost. Motivated by this evidence base, the present study compares three representative CNN based segmentation architectures U-Net, LinkNet, and DeepLabV3+ for annual solar farm mapping from Sentinel-2 seasonal composites in Ba Ria–Vung Tau.

2.2 Multi-temporal monitoring of land-cover and infrastructure change

Multi-temporal land-cover monitoring and change detection are well-established topics in remote sensing, widely applied to urban growth, agricultural dynamics, and forest loss. Recent studies increasingly use dense optical time series, such as Sentinel-1/2 and Landsat-8/9, together with machine learning or deep learning to generate temporally consistent maps and quantify land-cover transitions over time [12, 10]. Temporal compositing and time-series feature extraction are par-

ticularly effective in tropical regions, where persistent cloud cover often limits single-date analysis [10].

For solar energy infrastructure, early studies mainly produced static inventories or single year maps. More recent work has developed temporally explicit solar datasets by applying deep learning segmentation to multi-year imagery, enabling the detection of newly built, expanded, and decommissioned solar facilities at large scales [7, 8]. Related frameworks have also shown the value of using multi-year image stacks or spectral indices to capture change trajectories rather than mapping each year independently [11]. Following this direction, our study uses annual Sentinel-2 composites and post-segmentation differencing to map new, persistent, and expanded solar farm areas at provincial scale.

2.3 Solar energy and remote sensing in Vietnam

In Vietnam, the use of remote sensing and AI for energy and environmental applications has grown rapidly in recent years, although studies specifically targeting solar farm monitoring remain limited. Sentinel-2 time series and CNN-based approaches have been increasingly applied to produce high-resolution land use and land cover maps, including products designed specifically for Vietnam and optimized for cloud-prone conditions through multi-year compositing [10]. Similar remote sensing and machine learning methods have also been used for forest monitoring, urban expansion, surface temperature analysis, and seawater quality assessment, highlighting their practical value for environmental management in Vietnam [13–15].

Although Vietnam has experienced rapid growth in both utility-scale and rooftop solar deployment, deep learning-based monitoring of solar PV infrastructure at provincial scale is still scarce. Existing applied studies have mainly focused on rooftop solar mapping in selected provinces and specific years, with limited attention to consistent annual monitoring of ground-mounted solar farms [16]. This study addresses that gap for Ba Ria–Vung Tau and demonstrates the potential of medium-resolution imagery and CNN-based segmentation for reproducible year-to-year monitoring of solar PV expansion.

3 Study area and data

3.1 Study area

Ba Ria–Vung Tau (BRVT) province, covering approximately 1,980 km^2 in southeastern Vietnam, is a strategically important region characterized by a diverse landscape of industrial zones and coastal lowlands. Due to its tropical monsoon climate, the province benefits from high annual solar irradiance year-round. This favorable resource, combined with available flat terrain near existing grid infrastructure, has accelerated the deployment of utility-scale solar farms since 2019. The geographic extent of the study area and the spatial distribution of these solar energy zones are illustrated in Fig. 1.

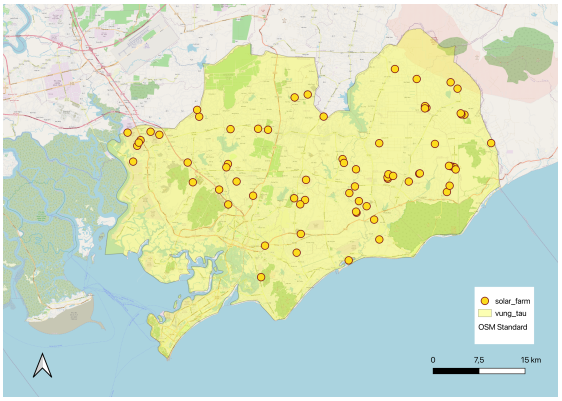

Fig. 1. Study area: Ba Ria–Vung Tau province, Vietnam. The base map shows the provincial boundary, locations of known solar farm clusters.

3.2 Sentinel-2 data

Sentinel-2 provides multispectral optical imagery in the visible, near-infrared (NIR), and shortwave infrared (SWIR) regions at 10–20 m spatial resolution, with a revisit frequency suitable for annual compositing in tropical environments [17]. These characteristics make it well suited for provincial-scale solar farm mapping and monitoring.

In this study, Sentinel-2 Level-2A surface reflectance data were acquired for Ba Ria–Vung Tau province over the period 2020–2025 using Google Earth Engine [18] and the `COPERNICUS/S2_SR_HARMONIZED` collection. Data from 2020–2022 were used to support annotation and model development, while imagery from 2023–2025 was used for annual solar farm mapping and interannual change analysis. This temporal separation helps ensure that the monitoring stage is conducted on imagery distinct from that used for label preparation. Table 1 summarizes the spectral inputs used in this study.

Table 1. Sentinel-2 bands and PVSI used for solar farm mapping.

Band / Index	Name / Description	Role in Analysis
B2	Blue	RGB composite / visualization
B3	Green	RGB composite / visualization
B4	Red	Vegetation contrast
B6	Red Edge 1	Vegetation / panel detection
B8	NIR	Vegetation / open ground
B11	SWIR-1	Panel / built-up discrimination
B12	SWIR-2	Surface material contrast
PVSI	Photovoltaic Spectral Index	Solar panel highlight

4 Methodology

4.1 Overall framework

The proposed research framework consists of four sequential phases: (1) data acquisition and preprocessing, including Sentinel-2 image retrieval via Google Earth Engine (GEE), cloud masking, annual compositing, spectral feature construction, and manual labeling; (2) deep learning model development, involving the implementation and training of U-Net and LinkNet segmentation architectures; (3) model evaluation and selection, employing standard segmentation metrics to compare model performance and identify the most suitable architecture; and (4) annual solar PV mapping and change monitoring, which generates yearly solar farm footprint maps and analyzes spatiotemporal expansion patterns. Figure 2 illustrates the overall processing chain.

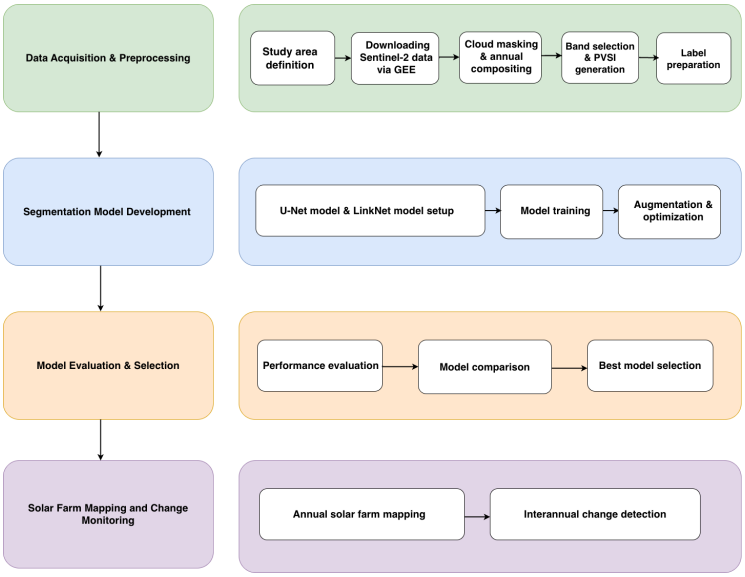

Fig. 2. Overall framework for annual solar farm monitoring using multi-temporal Sentinel-2 imagery and deep semantic segmentation. The pipeline proceeds from data acquisition and preprocessing through model training and comparison, annual binary map generation, and interannual change analysis.

4.2 Sentinel-2 data preparation

Study area gridding and annual image collection to support province-wide processing, Ba Ria–Vung Tau was divided into regular grid cells of approximately $0.25° \times 0.25°$, and only cells intersecting the provincial boundary

were retained. This grid-based strategy enabled systematic image extraction and reduced the computational burden of large-area processing in Google Earth Engine.

Sentinel-2 Level-2A surface reflectance imagery was collected for all years from 2020 to 2025 within a fixed acquisition window from January 1 to July 1. Annual composites from 2020–2022 were used for annotation and model development, whereas composites from 2023–2025 were reserved for province-wide inference and change monitoring. This design ensures a clear temporal separation between training data preparation and operational deployment.

Cloud masking and annual compositing To reduce cloud contamination, each Sentinel-2 image was paired with the corresponding cloud probability product and masked using a cloud probability threshold of 50%. All valid observations within each year were then aggregated using median compositing to generate a single annual cloud-reduced image. This approach suppresses residual cloud artefacts and produces spectrally stable annual representations suitable for interannual comparison.

Band selection and PVSI construction Each annual composite retained seven Sentinel-2 bands: B2, B3, B4, B6, B8, B11, and B12. In addition, a photovoltaic spectral index was computed as an auxiliary feature to enhance the discrimination of solar panel surfaces from surrounding land cover types. Following [19], PVSI was calculated as:

$$\mathrm{PVSI} = \frac{2.3B11 - 1.1B12 - B8}{2.3B11 + 1.1B12 + B8} + 0.5(B2 - B4 - B8) + \mathrm{sign}\left(1.3 - \frac{B6}{B8}\right) - 1 \quad (1)$$

where the input bands were first converted to surface reflectance by dividing digital values by 10,000. The resulting PVSI layer was appended to the multispectral stack, producing a final eight-channel input tensor. To assess its contribution, an ablation experiment was also conducted using the same architecture without the PVSI channel.

4.3 Solar farm segmentation

Training data preparation Training labels were generated through manual visual interpretation of the 2020–2022 annual Sentinel-2 composites. Solar farm polygons were delineated based on their distinct spatial patterns and spectral signatures, with cross-validation across consecutive years to minimize annotation inconsistencies. Only clearly identifiable photovoltaic installations were retained for the supervised training set.

The annotated images were converted into binary segmentation masks. Input patches of size 256×256 pixels were extracted, each containing eight channels: seven Sentinel-2 spectral bands (B2, B3, B4, B6, B8, B11, B12) and one derived

PVSI layer. A total of 5,000 labeled patches were divided into training, validation, and test sets in an 8:1:1 ratio (4,000 / 500 / 500). Importantly, no patches from the 2023–2025 monitoring period were included in the training dataset to ensure temporal generalization.

Data augmentation and preprocessing Data augmentation was applied during training using the Albumentations library to enhance generalization and reduce overfitting. The pipeline included horizontal flipping ($p = 0.8$), vertical flipping ($p = 0.8$), and random resized cropping (scale 0.5–1.0, $p = 0.5$), while maintaining the final patch size of 256×256 pixels.

Validation patches were padded when necessary to ensure consistent input size. All image tensors were converted from channel-last to channel-first format prior to training to comply with PyTorch conventions.

Segmentation models Semantic segmentation was performed using three encoder–decoder architectures from the `segmentation_models_pytorch` library: U-Net, LinkNet and DeepLabV3+.

U-Net employs a symmetric encoder–decoder structure with skip connections that propagate fine-grained spatial features from the encoder to the decoder, enabling precise boundary delineation. LinkNet adopts a more compact design in which encoder feature maps are directly added to decoder outputs, reducing parameter count while preserving spatial detail. DeepLabV3+ extends the encoder–decoder paradigm with atrous spatial pyramid pooling (ASPP) in the encoder bottleneck, designed to capture multi-scale contextual information through parallel atrous convolutions at multiple dilation rates.

All three models were configured to accept eight-channel input tensors (seven Sentinel-2 spectral bands and the derived PVSI layer) and produce a single binary output channel corresponding to solar farm presence. For all architectures, an EfficientNet-B7 encoder pretrained on ImageNet was adopted to ensure a consistent feature extraction baseline across the comparison. A sigmoid activation function was applied at the output layer to generate per-pixel probabilities of solar farm occurrence.

Loss function, optimization, and evaluation Training employed binary focal loss, which is well-suited for highly imbalanced segmentation tasks where solar farm pixels constitute only a small fraction of the image area. The Adam optimizer was used with an initial learning rate of 5×10^{-4} and a batch size of 16. Model performance was evaluated using Intersection over Union (IoU), F1-score, overall accuracy, precision, and recall, with a decision threshold of 0.5 for binary predictions.

4.4 Annual mapping and post-processing

Following model selection, the LinkNet model, trained on 2020–2022 imagery, was applied to annual composites from 2023–2025. Because the monitoring-

period imagery was not used during training, this setup provides a form of prospective validation and reflects the model's operational generalization. Predicted probability maps were thresholded at 0.5 to generate binary solar farm masks, which were then mosaicked to the full provincial extent for each year.

To improve map quality, post-processing included: (1) removal of connected components smaller than 0.5 ha, consistent with the practical detection limit of utility-scale installations at Sentinel-2 resolution; (2) morphological opening to remove isolated noisy clusters; and (3) morphological closing to smooth minor boundary irregularities and improve spatial consistency across years.

The final output consists of three georeferenced binary raster maps (M_{2023} to M_{2025}), representing the estimated solar farm footprint for each year.

5 Experimental setup

Table 2 summarizes the key parameters of the experimental configuration. Data acquisition, cloud masking, and annual compositing were performed on Google Earth Engine. Exported GeoTIFF composites were subsequently tiled on local hardware and used for model training, evaluation, and inference.

Table 2. Experimental configuration for segmentation model training and inference.

Parameter	Value
Segmentation models	U-Net, LinkNet, DeepLabV3+ (PyTorch implementation)
Input channels	8 (B2, B3, B4, B6, B8, B11, B12, PVSI)
Tile/patch size	256×256 pixels at 10 m
Training / Val / Test	80% / 10% / 10%
Encoder backbone	EfficientNet-B7 (ImageNet pretrained)
Loss function	Binary focal loss
Optimizer	Adam
Initial learning rate	5×10^{-4}
Batch size	16
Compositing period	January–June per year
Annotation composites	2020, 2021, 2022
Monitoring composites	2023, 2024, 2025
Output activation	Sigmoid
Binarization threshold	0.5

Model performance was evaluated on the held-out test set using standard segmentation metrics, including **Precision**, **Recall**, **F1-score**, **Intersection over Union (IoU)**, and **Overall Accuracy (OA)**. Precision measures the proportion of correctly predicted solar farm pixels among all predicted positives, while Recall quantifies the proportion of reference solar farm pixels successfully detected. F1-score summarizes the balance between precision and recall, IoU evaluates the spatial overlap between predicted and reference solar farm regions, and OA represents the proportion of correctly classified pixels across all classes.

6 Results and analysis

6.1 Segmentation performance

Table 3 summarizes the segmentation performance of the evaluated models on the held-out test set. Three architectures were assessed using the same 8-channel input configuration (seven Sentinel-2 spectral bands and PVSI): U-Net, LinkNet, and DeepLabV3+.

Table 3. Segmentation performance of the evaluated architectures on the held-out test set using the full 8-channel input configuration.

Model	Precision (%)	Recall (%)	F1 (%)	IoU (%)	OA (%)
U-Net (8-band + PVSI)	97.00	93.07	93.31	90.60	99.27
LinkNet (8-band + PVSI)	97.77	92.50	93.32	90.74	99.31
DeepLabV3+ (8-band + PVSI)	98.75	88.25	90.93	87.62	98.78

U-Net and LinkNet both achieved strong performance (IoU $>$ 90%, F1 $>$ 93%), confirming the effectiveness of the proposed multispectral input. LinkNet provided the best overall balance, with the highest IoU (90.74%), F1-score (93.32%), and OA (99.31%), while U-Net achieved slightly higher recall (93.07%). In contrast, DeepLabV3+ achieved the highest precision (98.75%) but lower recall, IoU, and F1-score, indicating more conservative predictions and greater omission of boundary and sparse-array pixels, which is less suitable for annual area estimation and change monitoring.

Overall, the results indicate that lightweight encoder–decoder architectures are more suitable for provincial-scale solar farm mapping from Sentinel-2 imagery. To further evaluate the contribution of PVSI, an ablation experiment was conducted using LinkNet (Table 4).

Table 4. Ablation study on the contribution of PVSI using the LinkNet architecture.

Input Configuration	Precision (%)	Recall (%)	F1 (%)	IoU (%)	OA (%)
LinkNet (7-band, no PVSI)	98.12	91.77	93.19	90.32	99.23
LinkNet (8-band, with PVSI)	97.77	92.50	93.32	90.74	99.31

Including PVSI consistently improved recall, F1-score, IoU, and OA, with only a slight reduction in precision. In particular, recall increased by 0.73 percentage points and IoU by 0.42 percentage points, indicating that PVSI helped reduce omission errors and improve detection completeness. Although the gain was moderate, it is meaningful for multi-year monitoring, where small per-pixel improvements can contribute to more reliable annual area estimates and change trajectories.

Visual inspection further confirmed that U-Net and LinkNet reliably delineated the main extent of large and compact solar installations, with only minor boundary uncertainty in fragmented or irregular arrays. DeepLabV3+ generally produced cleaner but more spatially conservative predictions, often undersegmenting peripheral pixels. Representative examples are shown in Figure 3.

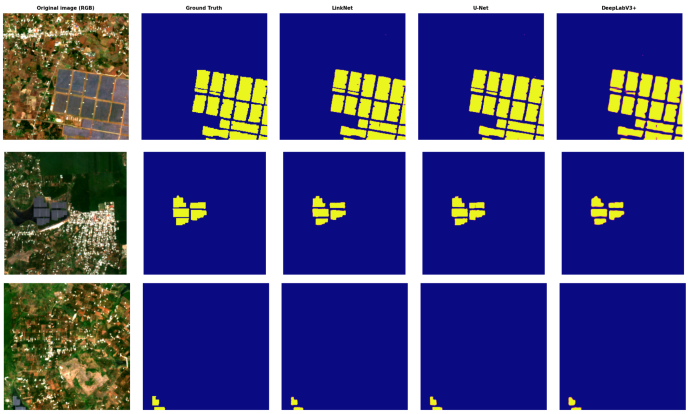

Fig. 3. Representative segmentation examples. Each row shows (left to right): original Sentinel-2 image, reference annotation, LinkNet prediction, U-Net prediction and DeepLabV3+ prediction.

6.2 Annual solar farm mapping

Figure 4 presents the annual solar farm footprint maps for 2023–2025.

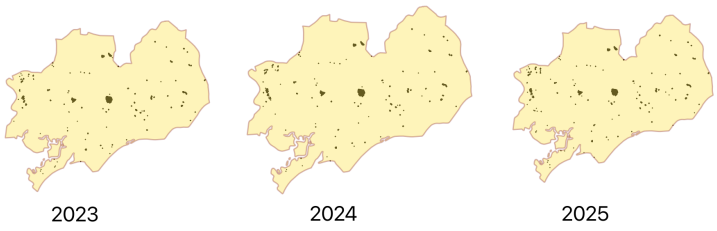

Fig. 4. Annual solar farm maps for Ba Ria–Vung Tau, 2023–2025.

The spatial distribution of solar farm installations from 2023 to 2025 reveals a clear pattern of gradual expansion and stabilization. A noticeable increase in the number and density of solar farm pixels is observed between 2023 and

2024, indicating a phase of active infrastructure development. In contrast, the transition from 2024 to 2025 shows a slower growth rate, with changes primarily characterized by the densification and minor expansion of existing clusters rather than the emergence of large new installations.

Spatially, solar farms exhibit a clustered distribution pattern, with several core regions maintaining persistent occupancy across all three years. These core clusters expand outward over time, while smaller satellite installations appear sporadically in surrounding areas. This pattern suggests a transition from rapid expansion to a more stable and mature development phase of solar energy infrastructure in the study area.

6.3 Change monitoring from 2023 to 2025

Table 5 summarizes the annual solar farm area estimates and derived interannual change statistics for the available monitoring period.

Table 5. Annual solar farm area estimates and interannual net change in Ba Ria–Vung Tau from 2023 to 2025.

Year	Total Area (ha)	Net Change (ha)	Cumulative Growth (%)
2023	413.18	—	—
2024	415.63	+2.45	+0.59
2025	418.64	+3.00	+1.32

The change monitoring results indicate a modest but consistent increase in mapped solar farm extent over 2023–2025. Total estimated area grew from 413.18 ha in 2023 to 418.64 ha in 2025, representing a cumulative increase of 1.32%. Annual net gains remained limited, with 2.45 ha added between 2023–2024 and 3.00 ha between 2024–2025.

Spatially, most changes were associated with incremental expansion around existing solar farm clusters rather than the appearance of entirely new large-scale sites. The overall spatial pattern remained highly stable across the three years, indicating that the proposed framework provides reliable annual monitoring at the provincial scale. Apparent loss areas were minimal and are more likely attributable to boundary-level classification variability than actual infrastructure removal.

7 Discussion

7.1 Value of Sentinel-2 for regional solar farm monitoring

The results confirm that Sentinel-2 provides a practical basis for provincial-scale solar farm monitoring in Ba Ria–Vung Tau. Although its spatial resolution is

moderate, medium- and large-scale solar installations can be mapped consistently when annual compositing and suitable spectral inputs are used. In particular, the combination of multispectral Sentinel-2 bands and PVSI improved the separability of solar farm surfaces from surrounding land cover.

A further advantage of Sentinel-2 is its frequent revisit cycle, which supports cloud-reduced annual compositing under tropical conditions. Given its free availability and operational continuity, Sentinel-2 is well suited for repeatable and cost-effective long-term monitoring.

7.2 Comparative performance of segmentation architectures

The comparative experiments show that U-Net and LinkNet are both highly effective for medium-resolution solar farm segmentation, with IoU values above 90% and F1-scores above 93%. Their encoder–decoder structure with skip connections appears well suited to the spatial regularity of utility-scale solar farms. LinkNet achieved the best overall balance, while U-Net showed slightly higher sensitivity in recovering solar farm pixels.

By contrast, DeepLabV3+ produced the highest precision but lower recall and weaker overall spatial agreement. This more conservative prediction behavior is less desirable for monitoring applications, where omission errors can lead to underestimation of annual area and expansion. The PVSI ablation experiment further confirmed that PVSI contributes useful spectral information, particularly by improving recall and overall detection completeness.

7.3 Monitoring perspective

A key strength of the proposed framework is its monitoring-oriented design. Rather than focusing only on single-date detection, the framework enables annual tracking of solar farm persistence, expansion, and spatial stabilization. The 2023–2025 results indicate that solar development in the study area was dominated by incremental expansion and densification of existing clusters rather than abrupt emergence of many new large-scale sites.

This temporal perspective is particularly valuable for infrastructure monitoring and regional planning, where year-to-year spatial evolution is often more informative than a single static inventory.

7.4 Limitations and future directions

Several limitations should be acknowledged. First, Sentinel-2 at 10–20 m resolution is appropriate for utility-scale solar farms but not for reliable detection of rooftop or very small distributed installations. Second, some spectral ambiguity remains possible in complex built-up areas despite the use of SWIR bands and PVSI. Third, annual compositing improves cloud robustness but may introduce slight temporal lag for installations constructed near the end of the compositing window.

In addition, the training labels were manually prepared and therefore remain dependent on expert interpretation. Future work could improve spatial detail through higher-resolution optical imagery, increase robustness through Sentinel-1 SAR integration, and further enhance temporal analysis using spatiotemporal deep learning or uncertainty-aware segmentation.

8 Conclusion

This study presented a practical deep learning framework for annual solar farm mapping and short-term spatiotemporal monitoring in Ba Ria–Vung Tau province, Vietnam, using multi-temporal Sentinel-2 imagery. Annual cloud-reduced composites were generated and used as input to semantic segmentation models trained on manually interpreted reference data. U-Net and LinkNet were evaluated using an eight-channel input composed of selected Sentinel-2 bands and the photovoltaic spectral index.

Both models achieved strong segmentation performance, with IoU values above 90% and F1-scores above 93% on the held-out test set, confirming the effectiveness of the proposed input configuration and modeling strategy for provincial-scale solar farm mapping. Among the evaluated models, LinkNet provided the best overall balance of accuracy and spatial consistency and was therefore selected for annual mapping and change analysis.

The annual maps for 2023–2025 indicate gradual solar farm growth in the study area, driven mainly by localized expansion and densification of existing clusters rather than widespread emergence of new large-scale sites. These results demonstrate that medium-resolution optical imagery, combined with deep semantic segmentation and annual compositing, can provide a reliable and reproducible basis for operational solar infrastructure monitoring.

Overall, the proposed framework is cost-effective, repeatable, and transferable to other provinces in Vietnam or comparable tropical regions. Future work will focus on extending the monitoring period, integrating SAR and higher-resolution imagery, and exploring spatiotemporal deep learning approaches to further improve detection robustness and change characterization.

References

1. International Renewable Energy Agency: Renewable power generation costs in 2024 (2025), https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2025/Jul/IRENA_TEC_RPGC_in_2024_2025.pdf
2. Anh, P., Ngan, N.T.K., Huong, N.T.T.: Vietnam's national energy development strategy to 2030 and outlook to 2045. International Journal of Economics & Business Administration (IJEBA) 8(4), 1023–1032 (2020), https://ideas.repec.org/a/ers/ijebaa/vviiiy2020i4p1023-1032.html
3. Zhu, X.X., Tuia, D., Mou, L., Xia, G.S., Zhang, L., Xu, F., Fraundorfer, F.: Deep learning in remote sensing: A comprehensive review and list of resources. IEEE Geoscience and Remote Sensing Magazine 5(4), 8–36 (Dec 2017), http://dx.doi.org/10.1109/MGRS.2017.2762307

4. Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., Jiang, Y., Xu, H., Tan, W., Yang, Q., Wang, J., Gao, J., Zhang, L.: Deep learning in environmental remote sensing: Achievements and challenges. Remote Sensing of Environment (2020), https://api.semanticscholar.org/CorpusID:213463573
5. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation (2015), https://arxiv.org/abs/1505.04597
6. Kruitwagen, L., Story, K.T., Friedrich, J., Byers, L., Skillman, S., Hepburn, C.: A global inventory of photovoltaic solar energy generating units. Nature 598(7882), 604–610 (2021), https://doi.org/10.1038/s41586-021-03957-7
7. Robinson, C., Ortiz, A., Kim, A., Dodhia, R., Zolli, A., Nagaraju, S.K., Oakleaf, J., Kiesecker, J., Ferres, J.M.L.: Global renewables watch: A temporal dataset of solar and wind energy derived from satellite imagery (2025), https://arxiv.org/abs/2503.14860
8. Li, A., Liu, L., Li, S., et al.: Global photovoltaic solar panel dataset from 2019 to 2022. Scientific Data 12, 637 (2025), https://doi.org/10.1038/s41597-025-04985-y
9. Viennois, G., Tulet, H., Tresson, P., Ploton, P., Couteron, P., Barbier, N.: Sentinel-2 forest typology mapping in central africa: assessing deep learning and image preprocessing effects. Frontiers in Remote Sensing Volume 6 - 2025 (2025), https://www.frontiersin.org/journals/remote-sensing/articles/10.3389/frsen.2025.1682132
10. Truong, V.T., Hirayama, S., Phan, D.C., et al.: Jaxa's new high-resolution land use land cover map for vietnam using a time-feature convolutional neural network. Scientific Reports 14, 3926 (2024), https://doi.org/10.1038/s41598-024-54308-1
11. Cullerton, M., Zhu, Z., Qiu, S., Rittenhouse, C.D., Suh, J.W.: Back in time: A novel time series and deep learning framework for mapping solar installations. Science of Remote Sensing 12, 100322 (2025), https://www.sciencedirect.com/science/article/pii/S2666017225001282
12. Deep learning and remote sensing for agricultural land use monitoring: A spatio-multitemporal analysis of rice field conversion using optical satellite images. International Journal of Advances in Data and Information Systems 6(2), 290 307 (Jun 2025), https://ijadis.org/index.php/ijadis/article/view/1385
13. Hang, P.T.: Evaluating the performance of sentinel 2a and planetscope satellite imaging systems in forest cover classification at ta dung national park, dak nong, vietnam. VIETNAM JOURNAL OF FOREST SCIENCE (2) (Jun 2024), https://vjfs.vafs.gov.vn/js/article/view/888
14. Ngan, C.T.: Smart monitoring of water quality in ha long bay using remote sensing and artificial intelligence. VAST (2025), accessed 2026
15. Van, T.: Application of remote sensing gis to estimate the changes of surface temperature in ninh thuan. VNUHCM Journal of Environment and Earth Sciences (2024)
16. TransitionZero: Mapping vietnam's rooftop solar landscape with machine learning and sentinel-2. Technical report (2025), accessed 2026
17. European Space Agency: Copernicus Sentinel-2 (processed by ESA), MSI Level-2A BOA Reflectance Product, Collection 1. European Space Agency, https://doi.org/10.5270/S2_-znk9xsj, accessed: 2026-04-09
18. Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R.: Google earth engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment 202, 18–27 (2017), https://www.sciencedirect.com/science/article/pii/S0034425717302900
19. Shimada, S., Takeuchi, W.: A new spectral index to characterize solar photovoltaic panels for sentinel-2 data. In: Proceedings of the Asian Conference on Remote Sensing (ACRS 2022). Mongolia (2022)