EUROSUN2026
Enhancing Short-term Solar Irradiance Forecasting using Ground-based Pyranometers and Hemispherical Sky Imagers

Martin Laroux[1], Thomas Carrière[1], Nicolas Thevenin[1], Boilley Alexandre[1]
[1] Calibsun, 955 route des Lucioles, 06560 Sophia Antipolis, France
alexandre.boilley@calibsun.com

Summary

Tropical photovoltaic (PV) systems are severely impacted by small, fast-moving convective clouds that induce abrupt solar irradiance ramps, threatening grid stability, and reducing the efficiency and lifetime of coupled battery and hydrogen storage systems. Existing forecasting solutions based on numerical weather prediction and satellite data lack the spatial and temporal resolution required to anticipate these rapid fluctuations. Ground-based pyranometers provide accurate measurements, but cannot predict incoming cloud-induced irradiance drops. This study demonstrates the critical role of hemispherical sky-imagers for very short-term solar forecasting in tropical environments. Using a sky-imager coupled with a local pyranometer enables estimating cloud optical depth and its impact on ground solar irradiance. A deep learning algorithm is trained to convert hemispherical images onto irradiance on ground. This approach provides a robust framework for estimating irradiance ramps and represents a key step toward improving PV forecasting and enabling reliable integration of solar energy with advanced storage systems in tropical regions.
Keywords: solar irradiance, in-situ measurements, sky-imagers, deep learning, forecasting

1 Introduction

The rapid expansion of solar energy deployment is transforming power systems worldwide, with particularly strong growth in tropical regions where the solar resource is abundant. However, tropical climates are characterized by highly dynamic atmospheric conditions, dominated by small, fast-moving convective clouds that induce strong and rapid fluctuations in surface solar irradiance. These high-frequency variations pose significant challenges for the integration of photovoltaic energy (PV) into electrical grids, especially as energy systems increasingly rely on tight coupling between photovoltaic generation and energy storage technologies such as batteries and green hydrogen production. In this context, accurate very short-term forecasts, from a minute to tens of minutes ahead, are essential to optimize energy management, ensure grid stability, and maximize the efficiency of storage systems.

Although numerical weather prediction models and satellite observations provide valuable information at larger spatial and temporal scales, they often lack the resolution required to capture rapid local irradiance variability. Ground-based measurements therefore play a critical role in addressing this gap. Pyranometers provide precise, high-frequency observations of global horizontal irradiance, forming the backbone of local monitoring systems. In addition, hemispherical sky-imagers offer complementary information by continuously observing cloud cover, motion, and morphology. The combination of these instruments enables detecting and tracking of small-scale cloud structures responsible for short-term irradiance variability (Schmidt et al., 2016).

Recent studies uses deep learning methods of different types [1] to forecast Global Horizontal irradiance (GHI) or production directly from sky-imagers raw output. These methods outperform persistence forecast but lack in anticipating ramps. Cloud base height can be used to better estimate cloud shadow on ground [2]. It requires a ceilometer on site or to setup multiple cameras in the area for stereoscopic vision. A ceilometer is an expensive measurement system compared to sky-imagers and cannot be installed on all customers sites. Stereoscopic vision requires a very precise segmentation of clouds in the images to determine the position of each clouds on the different synchronized images. This segmentation and the identification of similar cloud between two cameras is a complex task in tropical dense small clouds conditions. [3] integrate physical information in the neural network to improve the ramp time alignment as well as peak errors. This demonstrates that neural networks require additional information to anticipate cloud movement or cloud formation and extension in convective conditions.

The aim of this study is to present a method using physical information from the last image from the sky-imager but also temporal information from the past images. However, when forecasting ground irradiance based on sky-imagers, one of the first step is to determine cloud position in the image to then compute their displacement and forecast irradiance drop. In the sun area, which is the area where clouds impact solar radiation received on ground at the location of the camera, pixel values are saturated and segmentation algorithm exclude the area.

In this area, small and fast moving clouds becomes saturated due to sun diffraction and tends to extend the excluded sun area and disappear from the segmented image. In this study we determine, in previous images which cloud will cross the direction of the sun. Based on its location (azimuth and elevation), its pixel color and its effect on ground measured irradiance when effectively obstructing the sun rays, a Gradient Boosting model is trained to reconstruct an irradiance map. This map can then be used as input for forecasting methods.

2 Dataset

The dataset this study is based on consists on 1 year and 3 months of 15s Mobotix Q71 sky images in West India with time align pyranometer measurement as well as 1 year in North of France.
Climate in India is Arid Hot subject to tropical monsoon periods between June and September. Only monsoon periods (2025-2026) are kept in this study and images where dirt is detected are removed. Data when one of the pyranometer or the sky-imager is missing were discarded from the dataset. Less than 6 months of images were used. Climate in North of France is Oceanic, only dirt periods and non coincident data were removed and almost all year is available.
Ground solar irradiance measurements have a 1 minute frequency and where resample to 5 min to limit high frequency, therefore images in the mean interval were removed to be timely aligned with measurements.

3 Methodology

To obtain a map of expected irradiance attenuation from the raw image, a series of expert models are trained and operated jointly. Their initial input are the raw outputs of the sky imagers, that are png images constituted of 3 channels red, green and blue (R, G, B).
First, a clear-sky model provides a theoretical picture of the sky if there were no clouds at all. This model gives the color of each pixel depending on its Solar Position Angle (SPA), that is its angular distance to the Sun, and its Polar Zenith Angle (PZA), that is the pixel aim direction elevation.
Then, we combine a raw colorimetric segmentation with a Global Cloud Context Model (GCCM) and theoretical clear-sky picture to provide a final cloud segmentation of the original picture. The GCCM model is a probabilistic classifier that forecasts the probability that the cloud context belongs to each of these three classes: clear-sky, mixed or overcast. The colorimetric mask is linearly combined with a full-cloud mask and full-sky mask using these probabilities to obtain the final cloud mask.
Finally we use an ground-irradiance model that is trained to locally estimate the expected cloud attenuation for any given subset of the picture. By rolling this model over the full picture we are able to generate an ground-irradiance map. In this study, we compared two possible models: a Gradient Boosting (GB) and a Convolutional Neural Network (CNN).
The general workflow to obtain the attenuation map is shown on figure 1. The training of the different components of the workflow is presented in the following sections.

3.1 Clear sky generation

The first required component is the clear-sky generation model that estimates the theoretical picture that the imager should see if there were no clouds. To do so, a collection of clear-sky pictures is first obtained by using fixed thresholds on colorimetric indicators on the pictures. These thresholds are set to obtain a large number of false positives e.g. mixed or overcast skies incorrectly classified as clear. The idea is to filter out a lot of pictures easy to dismiss, without taking the risk of reducing the variety of clear-skies in the dataset. The obtained dataset is then manually filtered to remove incorrectly classified pictures, which is facilitated by the use of this first filter. This process is summarize on figure 2
Then for each channel, a model of the clear-sky pixel color as a function of its SPA and elevation is fitted numerically as shown on figure 3. The points are binned into regular intervals, and the color estimator is fitted as a numerical interpolator between the bin medians.

3.2 Cloud segmentation

The cloud segmentation is obtained by combining a raw colorimetric segmentation with context information from the GCCM. This process is described in the following sections.

3.2.1 Raw segmentation

The raw segmentation is based on the Red/Blue ratio (RBR):

$$RBR = \frac{R}{B + \epsilon} \tag{1}$$

where ϵ is a small constant to avoid dividing by 0. A morphology opening is applied to pixels away from sun halo $SPA > 15°$, where $B > 0$ and $RBR < 0.75$. If more than 30% of sky is estimated as blue sky :

$$RBR_{CLS} = \frac{R_{CLS}}{B_{CLS} + \epsilon} \tag{2}$$

This factor is used to correct RBR based on light intensity:

$$RBR_{mod} = RBR - RBR_{CLS}(aS - b(I - 200)) \tag{3}$$

where a=0.5, b=0.005, S is the proportion of very bright pixels in the circum solar area. I is the perceptual luminance defined as follows:

$$I = 0.299R + 0.587G + 0.114B \tag{4}$$

An Otsu threshold is computed and limited to a minimal value of 0.39. If less than 30% of sky is estimated as blue RBR is kept as working variable and a fixed threshold of 0.75 is applied.

3.2.2 Global Cloud Context Model

To refine the raw segmentation, we train a Global Cloud Context Model (GCCM). This model is trained on a fully manually labeled set of images. This model returns 3 probabilities corresponding to each category as shown on figure 4. It uses 14 features based on averaged normalized colors and solar geometry, as presented in 3.2.2.

Feature name	Formula if applicable
Fraction of blue pixels	$mean(B) + 0.5.std(B)$
Variation coefficient	$\frac{std(B)}{mean(B)}$
Average RBR	$\frac{mean(R)}{mean(B)}$
Average normalized RBR	$\frac{R-B}{R+B}$
Contrast	$\frac{P_{90}-P_{10}}{mean(B)}$
Blue channel percentiles	$P_{10}, P_{25}, P_{50}, P_{75}, P_{90}$
Interquantile distance	-
Blue skewness	-
Local variance on a 8x8 subgrid	-
Solar elevation	-

The labeled sky-imager pictures are randomly split into a training and testing set. Four candidate models (Random Forest, Gradient Boosting, XGBoost and an ensemble) are trained on the training set using a 5-fold cross validation. A F1 score is computed for each category (clear sky, partially cloudy and overcast) and a macro F1 score is computed as the averaged of the 3 scores. Following the cross-validation, the best model is selected and trained on the overall train dataset.
The GCCM process an image and produce a probability for each of the following categories: clear sky, fully overcast or partially cloudy. This is used to mitigate a deterministic model. Therefore the final cloud classification map is a weighted average of a clear sky image with a 0 mask of cloud with a fully overcast mask of 1 and a cloud mask generated by the segmentation model presented above.

3.3 Ground-irradiance model

The last component of our process is the model that estimates the attenuation caused by a cloud as represent on a patch of the original picture. To do so, we first obtain a training dataset by backtracking a cloud for which we could measure the attenuation with a pyranometer. Then we train a regression model to estimate the attenuation based on the Region of Interest (ROI) containing the cloud. This process is summarized on figure 5

3.3.1 Backtracking of the cloud ROIs

Instead of computing the shadow projected on ground from sky-imager, we process images backward in time to determine the image zone (40 pixels x 40 pixels) which will pass in front of the sun and generate a drop in ground irradiance. From the instant of the measured drop, the DeepFlow method, a fast dense optical flow computation based on robust local optical flow (RLOF) algorithms and sparse-to-dense interpolation scheme, is applied. The RLOF is a fast local optical flow approach described in [4] similar to the pyramidal iterative Lucas-Kanade method [5]. The algorithm goes back in time up to 7 images (40 minutes) computing cloud displacement between each time step. The retained ROI is computed using the centroid of 4 estimations of the location :

- median: based on Otsu threshold segmentation a median value of only cloudy pixels. This method is robust in term of speed but is very sensitive to cloud segmentation
- global: a mean of all pixels speed, deep sky included
- non null: keep only the speed of moving pixels, remove speed < 3 pixels per frame.
- weighted: weighted mean of pixel speed. Weight is inversely proportional to distance of the estimated location of the cloud.

If the ROI is outside the image, if displacement speed exceeds 40 pixels or if the zone location given by all 4 methods is distant of more than 50 pixels the sample is rejected.
With this method, we produce a dataset for training the ground irradiance model. To reduce uncertainty the dataset is cleaned using physical method by comparing the clearness index K_c defined as the ratio of global horizontal irradiance (GHI) on clear sky global horizontal irradiance (GHI_{CLS}) :

$$K_c = \frac{GHI}{GHI_{CLS}} \tag{5}$$

The theoretical clear sky is obtained from CAMS McClear [6].
Besides, if the GCCM forecasts a full clear sky for the candidate picture but the K_c is low or if the GCCM forecasts a full overcast but the K_c is high, the samples are removed as well as if the K_c is below 0.1.
Finally the structural similarity index measure (SSIM) from the patch series used to backtrack a ROI. If this index is below 0.2 all the following patches are removed from the dataset.

3.3.2 Model definition and training

Using this dataset 2 types of model are trained: a Xgboost model and a CNN model. These two models are trained independently and their performances are estimated in the following section. Both models use scalar features consisting on solar position (elevation, azimuth), cyclical encoding of the calendar date, theoretical clear sky, fraction of cloud over the full image, the fraction of cloud within the patch, mean luminance of cloud pixels within the patch, class probabilities from GCCM. To rigorously evaluate the generalization performance while accounting for the strong spatiotemporal correlations in our all-sky imager dataset, we employed a Leave-One-Day-Out (LODO) cross-validation scheme. This approach systematically excludes all image patches from a single day for testing while training exclusively on data from the remaining days, thereby preventing any data leakage that would occur with random splits where temporally correlated samples from the same day could appear in both training and testing sets. The LODO validation exposed the true challenge of model generalization: days yielding negative R^2 values consistently corresponded to atypical meteorological conditions or underrepresented months (e.g., October/November) in the dataset. Given the severe class imbalance across months—with August containing 2400 patches compared to only 34 in November—conventional random validation splits produced artificially inflated performance metrics (R^2 0.88), obscuring the model's inability to handle infrequent or extreme conditions. Thus, LODO emerges as the most reliable validation framework for assessing generalization in datasets with strong daily autocorrelation, providing a critical methodological safeguard against overoptimistic performance estimates.

Hybrid CNN We propose a hybrid deep learning model that jointly processes raw sky image patches (40x40 pixels) and 11 contextual features to predict the Clearness Index (Kc) at the patch level. The architecture combines a Convolutional Neural Network (CNN) image encoder with a shallow Multi-Layer Perceptron (MLP) feature branch, fusing both representations before the regression head. A schematic architecture is presented in Figure 6.

Xgboost Three experts Xgboost models were trained on the three categories defined by the GCCM model on the French dataset. The Indian dataset is limited in size and temporal coverage, therefore, we explore three ways to assess the cross-site generalization capability of our deep learning model. (1) a French-only model applied directly to India, which yielded negative R^2 scores across multiple regimes (clear, mixed) despite capturing the qualitative variation in irradiance, indicating systematic miscalibration due to site-specific optical/physical differences; (2) a 100% Indian model trained from scratch, which achieved strong internal performance (R^2 ¿ 0.9 in some regimes) but suffered from high variance and overfitting risks attributable to the limited dataset size and reduced diversity; and (3) a transfer learning (TL) approach, where the French model was fine-tuned using a small subset of Indian data. The TL strategy emerged as the optimal compromise, achieving performance comparable to the Indian scratch model while avoiding the need for large local datasets, and demonstrating robustness across multiple iterations (2025–2026), including scenarios with reweighted extreme Kc values (clear/overcast) to prevent bias toward mixed conditions. In practice, TL was implemented by retaining the XGBoost trees trained on French data and adding a limited number of new trees fine-tuned on Indian samples, rather than full retraining from scratch. This approach effectively mitigated site-specific biases while preserving generalization capacity. To counteract the imbalance in the target dataset, a sample weight scheme is applied:

$$w_i = \begin{cases} 3 \quad if \quad K_c < 0.2 \quad or \quad K_c > 0.8 \\ 1 \qquad otherwise \end{cases} \tag{6}$$

This emphasizes the physically challenging clear and overcast extremes of the Kc distribution, which are underrepresented in the 2026 dataset. The GCCM model trained on french dataset was tested on a dataset of 1000 annotated Indian images and performed well. There was no need for transfer learning for this model.
Instead of direct image input as in the CNN model, several scalar features are extracted from an image: Blue and Red channels statistics, channel differences, center patch statistics, RBR and normalized RBR, normalized color statistics, normalized percentiles and gray level co-occurrence matrices (GLCMs) texture (33 features). Performance is evaluated on a held-out test set (15% of days, day-level split) using MAE and RMSE. Results are further broken down across three Kc bins to assess regime-specific performance.

4 Results and discussion

The quantitative evaluation of the French-only model, transfer learning (TL) model, and CNN across sky regimes reveals distinct performance trade-offs, with clear advantages emerging for TL and CNN in most conditions. All results presented hereafter are summarized in Table 1 for MAE and Table 2 for RMSE. For clear-sky conditions, the TL model achieves the lowest MAE (0.029) and RMSE (0.039), outperforming the French-only model (MAE: 0.044; RMSE: 0.050) and matching the CNN's performance (MAE: 0.028; RMSE: 0.039). This suggests that TL effectively corrects site-specific biases in clear-sky irradiance estimation, where aerosol scattering and lens transmittance differences between France and India are most pronounced.
In mixed regimes, which are the most challenging due to dynamic cloud interactions with solar irradiance, the CNN demonstrates superior performance with an MAE of 0.113 and RMSE of 0.146—significantly outperforming both the French-only (MAE: 0.127; RMSE: 0.173) and TL models (MAE: 0.127; RMSE: 0.166). The TL model shows no improvement over the French-only baseline in this regime, highlighting a key limitation: while TL excels in correcting systematic biases (e.g., clear-sky calibration), it struggles to capture the spatial variability of mixed cloud fields without explicit architectural adaptations.
For overcast conditions, the French-only and TL models perform comparably well (MAE: 0.040; RMSE: 0.059–0.085), while the CNN exhibits higher errors (MAE: 0.062; RMSE: 0.084). This discrepancy likely stems from the CNN's sensitivity to fisheye distortions under uniformly overcast skies, where global irradiance patterns are smoother and less informative for convolutional feature extraction. The TL model's strong performance here (RMSE: 0.059) underscores its robustness in regimes dominated by large-scale cloud cover, where pre-trained features from the French dataset remain highly transferable.
Transfer learning as a robust baseline: TL consistently reduces errors in clear-sky and overcast regimes, proving its value as a lightweight adaptation strategy for cross-site deployment. Its inability to improve mixed-regime performance, however, signals the need for regime-specific architectural adjustments or data augmentation (e.g., synthetic mixed-cloud patches).
CNN's mixed-regime superiority: The CNN's ability to outperform both baselines in mixed conditions (MAE reduction of 11% vs. TL) aligns with its capacity to model local cloud interactions. However, its weaker performance in overcast regimes suggests a trade-off between spatial sensitivity and generalization to uniform irradiance fields.

Dataset constraints: The limited size of the Indian dataset and its imbalance across regimes (e.g., only 125 clear-sky samples in the validation dataset compared to more than 2000 samples for mixed regimes) likely amplify these differences. For instance, the CNN's overcast errors may reflect insufficient training data for this regime, whereas TL's stability benefits from leveraging the larger French dataset.
Future directions: Hybrid approaches combining TL with CNN architectures (e.g., fine-tuning a CNN pre-trained on French data with Indian samples) could harness the strengths of both methods. Additionally, physics-informed loss functions (e.g., enforcing energy balance constraints) may mitigate regime-specific biases, particularly in mixed conditions where empirical models struggle.
These results emphasize that no single method dominates across all regimes, and the optimal choice depends on the target application's priorities—whether prioritizing clear-sky accuracy (TL), mixed-regime adaptability (CNN), or operational simplicity (TL). Further validation under Leave-One-Day-Out (LODO) cross-validation would be critical to confirm these trends across unseen temporal conditions.
Future work should prioritize expanding the Indian dataset with balanced sampling across regimes and exploring CNN architectures tailored to fisheye imagery, while validating performance under LODO to ensure true generalization.

	French-model	TL model	CNN
Clear-sky	0.044	0.029	0.028
Mixed	0.127	0.127	0.113
Overcast	0.040	0.040	0.062

Table 1: MAE comparison across different models and sky conditions

	French-model	TL model	CNN
Clear-sky	0.050	0.039	0.039
Mixed	0.173	0.166	0.146
Overcast	0.085	0.059	0.084

Table 2: RMSE comparison across different models and sky conditions

5 Conclusion

This study presented a new deep learning framework for estimating solar irradiance from hemispherical sky images. The method explicitly leverages upstream cloud motion information from hemispherical sky images and links it to co-located pyranometer measurements. By combining cloud context classification, cloud-region backtracking, and patch-level attenuation estimation, the proposed workflow captures both spatial cloud structure and its radiometric impact on ground measurements.
Several deep learning modeling strategies were assessed. Using a large France based dataset the model was able to achieve robust performance across different sky conditions. This model was not immediatly transferable to other locations, and its performance varied depending on the specific sky condition. Different strategies were explored to address this issue, including a French-only baseline, transfer learning, and a hybrid CNN approach. The results are overall positive and confirm the relevance of coupling image-based features with local irradiance observations. No single method consistently dominates the others under all sky conditions, highlighting complementary strengths across regimes. Future work will focus on improving each model, developing cross-model or hybrid usage strategies, extending validation and adaptation to a third site, and integrating the approach into an operational forecasting system to better anticipate irradiance drops.

6 Figures

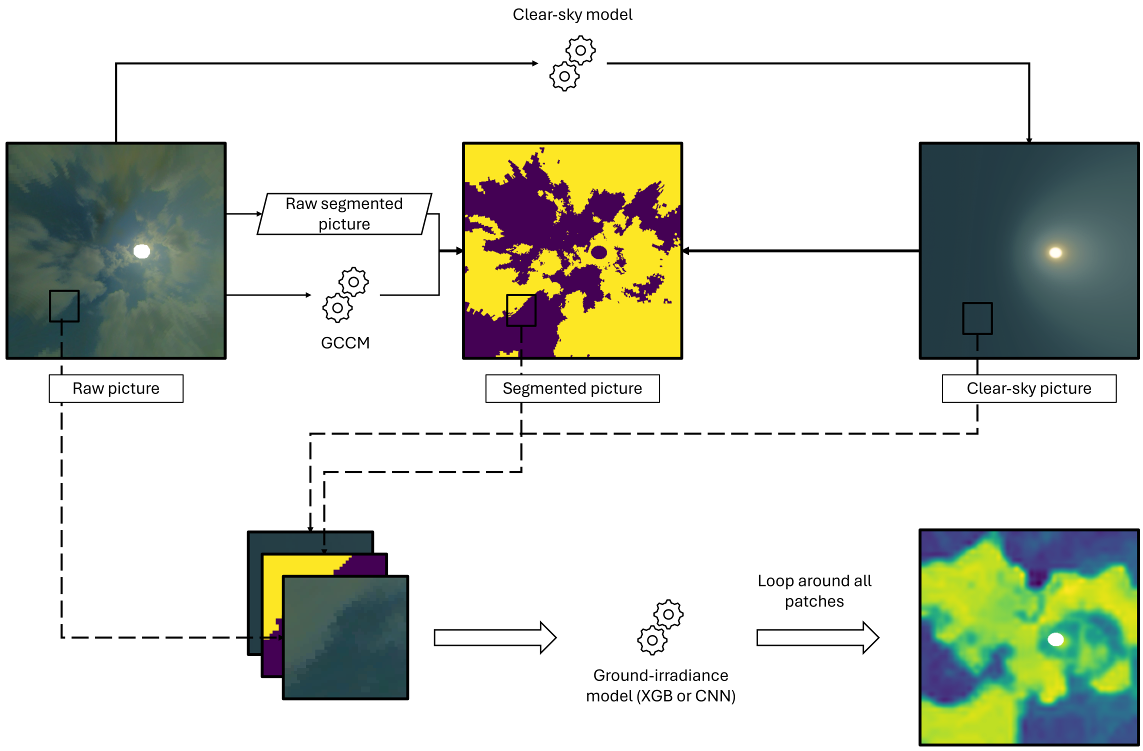

Figure 1: General workflow of the image to attenuation map. From the raw image, a cloud segmentation and a theoretical picture are obtained using two separate models. Then, a third model forecasts the local cloud attenuation on patches of the picture. The model is rolled over the whole picture to obtain a map.

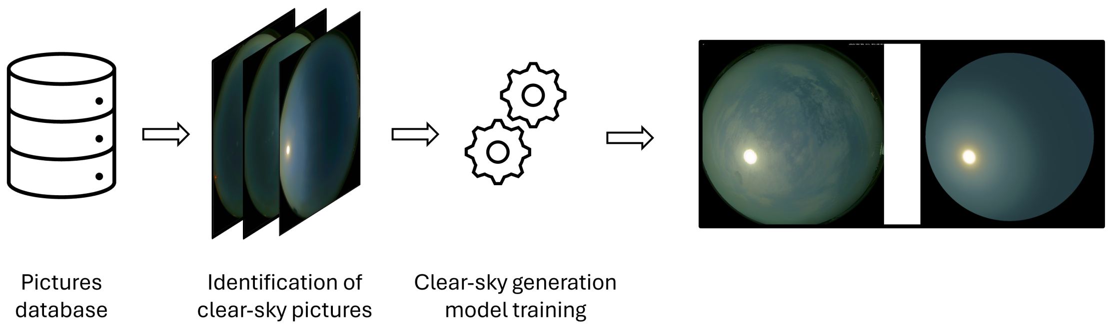

Figure 2: Process for training the clear-sky generation model

7 References

References

[1] Q. Paletta, G. Arbod, J. Lasenby, Benchmarking of deep learning irradiance forecasting models from sky images – an in-depth analysis, Solar Energy 224 (2021) 855–867. doi:https://doi.org/10.1016/j.solener.2021.05.056.
URL `https://www.sciencedirect.com/science/article/pii/S0038092X21004266`

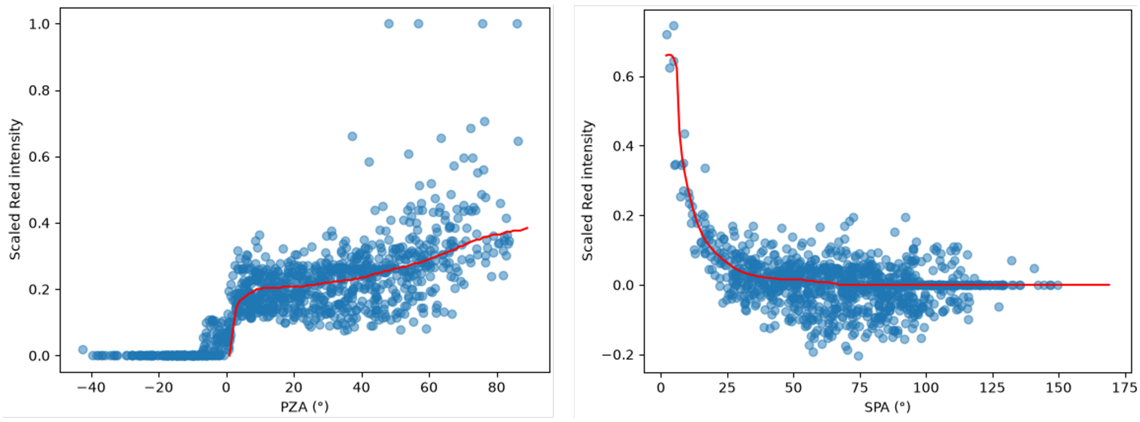

Figure 3: Scatterplot between pixel red color and PZA (left), SPA (right)

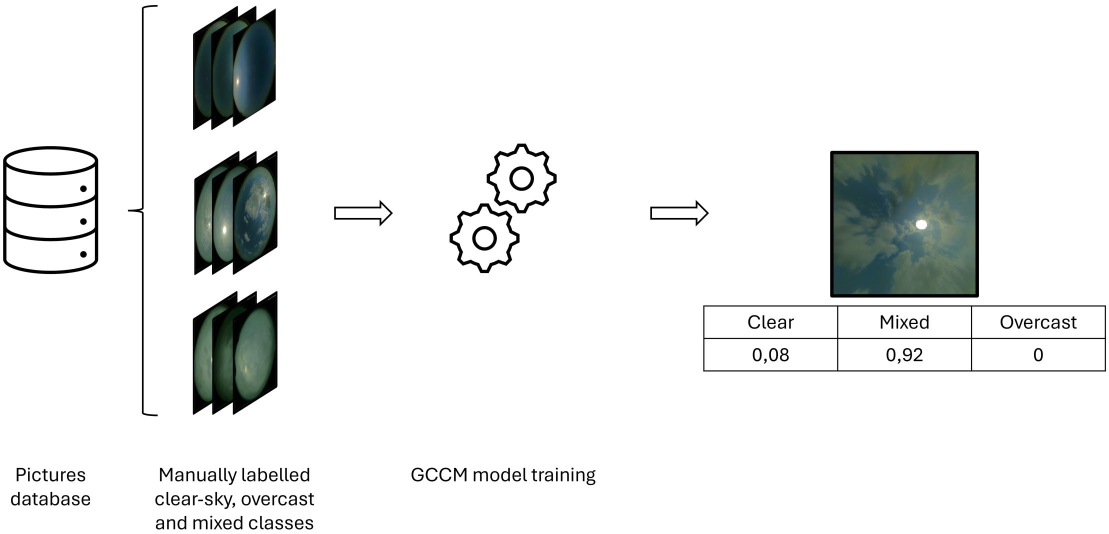

Figure 4: Training process of the GCCM model

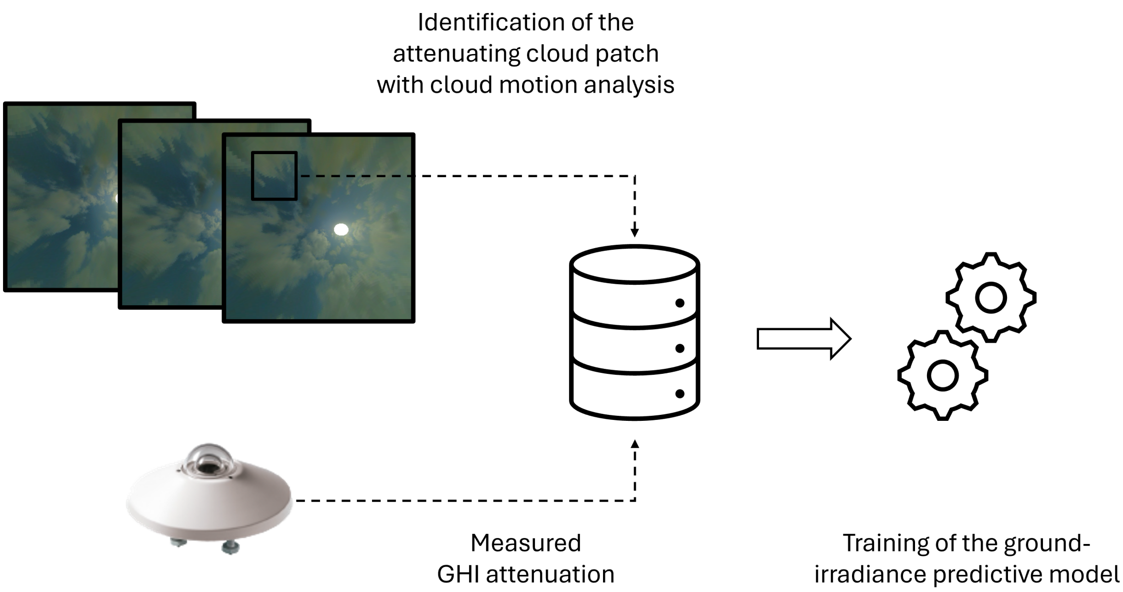

Figure 5: Training process of the ground-irradiance model

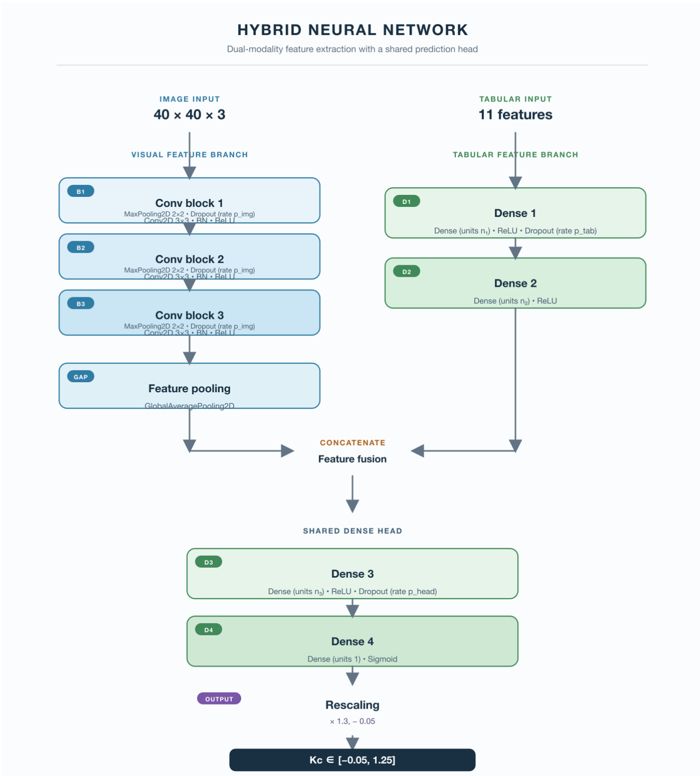

Figure 6: Hybrid CNN architecture

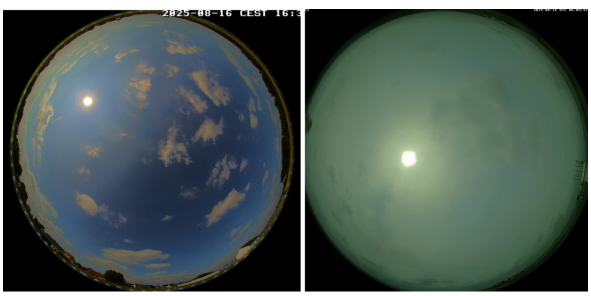

Figure 7: Image example from the North of France dataset (left). Image example from the Indian dataset (right).

[2] N. Straub, W. Herzberg, A. Dittmann, E. Lorenz, Blending of a novel all sky imager model with persistence and a satellite based model for high-resolution irradiance nowcasting, Solar Energy 269 (2024) 112319. doi:https://doi.org/10.1016/j.solener.2024.112319.
URL `https://www.sciencedirect.com/science/article/pii/S0038092X24000136`

[3] Y. Huang, X. Wang, W. Yao, X. Kong, H. Fan, Ultra-short-term pv power forecasting under complex sky conditions using all-sky imaging and physically informed dynamic features, Renewable Energy 274 (2026) 126183. doi:https://doi.org/10.1016/j.renene.2026.126183.
URL `https://www.sciencedirect.com/science/article/pii/S0960148126010098`

[4] T. Senst, V. Eiselein, T. Sikora, Robust local optical flow for feature tracking, in: IEEE Transactions on Circuits and Systems for Video Technology, Vol. 22, 2012, pp. 1377–1387.

[5] B. D. Lucas, T. Kanade, An iterative image registration technique with an application to stereo vision, Proceedings of Imaging Understanding Workshop (1981) 121–130.

[6] M. Lefèvre, A. Oumbe, P. Blanc, B. Espinar, B. Gschwind, Z. Qu, L. Wald, M. Schroedter-Homscheidt, C. Hoyer-Klick, A. Arola, A. Benedetti, J. W. Kaiser, , J.-J. Morcrette, Mcclear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions, Atmospheric MEasurement techniques 6 (2013) 403–2418. doi:doi:10.5194/amt-6-2403-2013.