CME LIVE OPEN ACCESS

Explorations in Astronomy & Astrophysics

Research Article

Vision-Based CNN Prediction of Sunspot Numbers From SDO/HMI Images

Fabian C. Quintero-Pareja , Diederik A. Montano-Burbano, Santiago Quintero-Pareja, and D. Sierra-Porta*

School of Digital Transformation, Technological University of Bolívar, Carlos Vélez Pombo Industrial and Technological Park, Km 1 Via Turbaco, Cartagena de Indias, Colombia.

Corresponding Author: D. Sierra-Porta, School of Digital Transformation, Technological University of Bolívar, Carlos Vélez Pombo Industrial and Technological Park, Km 1 Via Turbaco, Cartagena de Indias, Colombia.

Received: 2026 Feb 26 **Accepted:** 2026 Mar 20 **Published:** 2026 Mar 30

Abstract

Sunspot numbers constitute the longest and most widely used record of solar activity, with direct implications for space weather forecasting and heliophysical research. Traditional sunspot counting relies on visual inspection or algorithmic feature detection, both of which are limited by subjectivity, image quality, and methodological inconsistencies. Recent advances in deep learning, particularly convolutional neural networks (CNNs), enable the direct use of solar imagery for automated prediction tasks, reducing reliance on manual feature engineering. In this work, we present a supervised vision-based regression framework to estimate daily sunspot numbers from full-disk continuum images acquired by the Heli seismic and Magnetic Imager (HMI) onboard NASA's Solar Dynamics Observatory (SDO). Images from 2011–2024 were paired with daily sunspot numbers from the SILSO Version 2.0 dataset of the Royal Observatory of Belgium. After preprocessing and augmentation, a CNN was trained to predict scalar sunspot counts directly from pixel data. The proposed model achieved strong predictive performance, with R2 = 0.986 and RMSE = 6.25 on the test set, indicating close agreement with SILSO reference values. Comparative evaluation against prior studies shows that our approach performs competitively with, and in several cases outperforms, statistical and hybrid machine learning methods, while offering the novel advantage of bypassing explicit detection and manual feature extraction. Interpretability analyses using Grad-CAM and Integrated Gradients confirmed that the network consistently attends to sunspot regions when forming predictions. These results highlight the potential of deep vision-based approaches for operational solar monitoring, providing a scalable and automated pathway for real-time estimation of classical heliophysical indices. Future work should explore multimodal integration with magnetic field data and standardized benchmarking across solar cycles to strengthen operational applicability.

Keywords: Sunspots, Solar Cycle, Space Weather, Convolutional Neural Networks (CNNS), Deep Learning, Image Based Regression, SDO/HMI, SILSO V2.0

1. Introduction

Sunspots represent a fundamental and enduring measure of solar activity, with sunspot numbers providing the longest continuous record of solar variability, spanning more than 400 years [1,2]. The sunspot number is one of the most widely used indices in solar-terrestrial research due to its strong correlations with multiple space weather phenomena [3]. For example, the frequency and energy of solar flares, as well as the rate of coronal mass ejections, are well correlated with sunspot numbers, while cosmic ray fluxes exhibit an anticorrelated pattern over the solar cycle [3,4]. Sunspot observation is one of the oldest continuous scientific endeavors, with telescopic data spanning centuries and compiled into well-calibrated time series that reflect solar magnetic activity [1,2,5,6]. The foundation of modern sunspot quantification lies in the relative sunspot number, defined as k(10g +f), where f represents the number of individual sunspots, g the number of sunspot groups, and k a correction factor accounting for observer and instrumental differences.

Traditional manual methods for sunspot detection suffer from human subjectivity and inconsistencies, while automated techniques based on spatial filtering may compromise image resolution or distort physical sunspot properties [7-9]. Measurement accuracy has been shown to depend heavily on both the method and image quality, with systematic discrepancies noted between photographic records and sunspot drawing archives [10]. Technological advances have facilitated the development of high-resolution observational tools such as CCD imaging, spectroscopy, satellite platforms, and radio diagnostics, alongside tools for automated sunspot group area calculation, including software like S.C.A.T. (Solar Cycle Analyzer Tool) [11]. Recent efforts have leveraged image processing and deep learning techniques to improve automatic sunspot detection, although accurate grouping of

sunspots remains a challenge [2]. Modern computer vision methods, including Canny edge detection and topological analysis, have been applied to solar images from NASA's Solar Dynamics Observatory (SDO) for feature identification [12-14].

Recently, Acevedo proposed a machine learning method to predict sunspot numbers from solar imagery, leveraging topological feature extraction prior to modeling [15]. In contrast, our work employs fully automated convolutional neural networks to directly map full-disk solar images to sunspot number predictions, removing the need for handcrafted feature design and enabling end-to-end regression learning. The rise of deep learning especially Convolutional Neural Networks (CNNs) has revolutionized solar physics by enabling the analysis of massive, high-dimensional datasets. SDO alone generates over 1.5 terabytes of data daily, rendering traditional processing approaches insufficient [16]. Deep learning models can extract hierarchical spatial features directly from raw observational data without the need for handcrafted input representations [16-18]. CNNs have shown outstanding performance in characterizing solar active regions. Recognition accuracies above 95% have been reported for magnetic classification of sunspot groups, with Alpha-type accuracies reaching 98% and Beta-types above 88% [19,20].

Solar flare prediction is another key domain where CNNs have outperformed traditional methods. Recent models achieve True Skill Statistic (TSS) scores of 0.749 ± 0.079 for ≥M-class flares and 0.679 ± 0.045 for ≥C-class, surpassing earlier techniques [21,22]. Some architectures exceed TSS 0.80 for M-class prediction tasks [23,24]. Advances in attention-based deep learning have also enabled full-disk solar flare prediction, overcoming the limitations of early models that focused only on central disk regions (±30◦ to ±45◦) [21,22,25]. These models have demonstrated average recall rates near 0.52 for near-limb flares beyond ±70◦ [21]. The field now extends beyond classification, incorporating solar radio spectrum analysis where CNNs and LSTMs are used to identify and classify solar radio bursts across vast datasets [26,27]. Operational systems integrate ensemble learning, physics-informed models, and interpretable AI techniques such as attention maps and gradient attribution to enhance transparency [28]. The synergy between deep computer vision and solar physics big data has enabled real-time extraction of features from solar images with remarkable accuracy. Architectures like U-Net and ResUNet++ further refine segmentation performance for complex solar structures [29-31]. While deep learning has achieved impressive results in solar physics, several open challenges remain. Chief among these is the scarcity of high-quality, labeled datasets, as creating ground truth annotations demands expert effort and time. This scarcity is particularly acute for rare events such as X-class flares. Another limitation is model generalization: CNNs trained on specific datasets often struggle under changing observational conditions or across solar cycles. Instrumental variation, atmospheric distortion, and temporal evolution introduce domain shifts that hinder cross-cycle robustness.

Computational constraints also arise in real-time applications, where deep models may be too resource-intensive for timely forecasts. Additionally, interpretability remains an issue: most CNNs behave as black boxes, limiting the ability of researchers to trust or explain model outputs. Finally, inconsistent evaluation metrics and benchmarking datasets across the literature hinder fair comparison between methods. Most applications to date focus on classification rather than continuous regression tasks, leaving a gap in the direct prediction of quantitative indices such as sunspot numbers or magnetic field strength. In this study, we present a deep learning framework to estimate the daily sunspot number directly from full-disk digital images acquired by the Helioseismic and Magnetic Imager (HMI) aboard NASA's Solar Dynamics Observatory (SDO). We compile a dataset of continuum solar images from 2011 to 2024 and pair these with corresponding daily sunspot numbers from the SILSO Version 2.0 database maintained by the Royal Observatory of Belgium. Our approach leverages Convolutional Neural Network (CNN) architectures trained in a supervised regression setting, aiming to predict the continuous sunspot number as a function of the visual solar disk appearance alone. While prior research has extensively applied CNNs to tasks such as solar flare classification, sunspot group type recognition, or magnetic topology identification, few efforts have directly targeted the prediction of the scalar sunspot number from image data. This task poses distinct challenges, particularly due to the complex spatial distribution and varying contrast of sunspots across the solar disk, as well as the absence of public benchmarks for image-to- scalar sunspot regression. By addressing this gap, our work contributes a novel methodology that bridges solar image processing and quantitative sunspot analysis, expanding the scope of deep learning applications in operational space weather monitoring. Our results show that deep learning techniques are not only capable of modeling visual solar patterns but also hold potential for quantitatively bridging observational solar imagery and fundamental heliophysical indices. This work thus contributes to expanding the toolkit for space weather forecasting by integrating automated vision-based approaches into the prediction of classical solar metrics.

2. Data And Methods

2.1. Sunspot Number Data

Since 1981, the Solar Influences Data Analysis Center (SIDC) has served as the World Data Center for sunspot numbers, operating the SILSO (Sunspot Index and Long-term Solar Observations) project, which maintains the longest-running time series of solar activity measurements. The international sunspot number provided by SILSO is a vital reference for numerous scientific disciplines and supports over 500 institutional and individual users worldwide, including the International Astronomical Union and UNESCO. The revised SILSO Version 2.0 dataset, released in 2015, incorporates recalibrated sunspot counts with improved consistency across observatories and over time. This dataset continues

to rely primarily on visual hand-drawn sunspot counts from a network of ground-based stations, with the pilot station at Specola Solare Ticinese in Locarno, Switzer- land, serving as the primary reference [1]. The Version 2.0 revision corrected historical biases, homogenized the long-term series, and introduced an internally consistent standard that is now widely adopted in solar physics research. In this study, we used the daily total sunspot number covering the period from January 1st, 2011 to December 31st, 2024. This interval was selected to align with the availability of continuum solar images from the Heli seismic and Magnetic Imager (HMI) onboard NASA's Solar Dynamics Observatory (SDO). The daily sunspot number was extracted as a single scalar value representing the global level of solar magnetic activity, which served as the target variable for our supervised regression task. In addition to its long-term stability and scientific acceptance, this dataset is particularly well suited as a ground truth reference because it encodes the aggregate visual manifestation of sunspots across the entire solar disk, making it directly comparable to the image-based features learned by convolutional neural networks.

Figure 1 shows the daily sunspot number from SILSO (https://www.sidc.be/SILSO/datafiles) during the selected study period, including its smoothed profile that highlights the underlying 11-year solar cycle. The dataset spans the descending phase of Solar Cycle 24, the minimum around 2019, and the ascending phase of Solar Cycle 25 through 2024. This coverage ensures that the model is trained and evaluated across both low-activity and high-activity regimes, providing a more robust assessment of its generalization capacity.

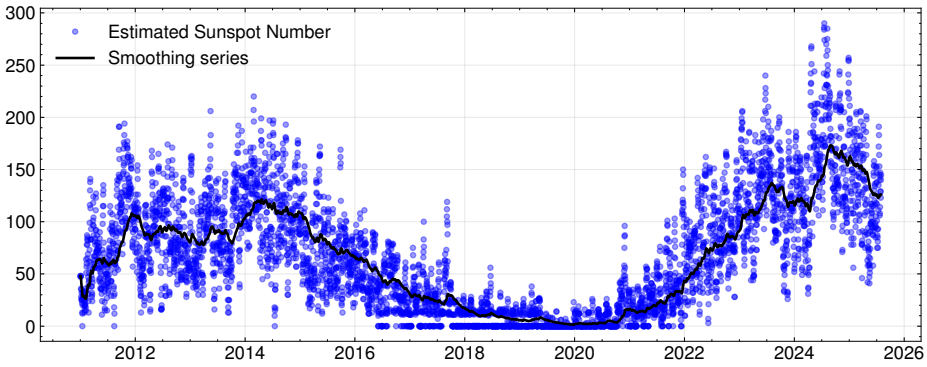

Figure 1: Daily Sunspot Number Time Series from SILSO (Version 2.0) between 2011 and 2024, Including the Smoothed Solar Cycle Trend (https://www.sidc.be/silso/datafiles). This Dataset Served as the Reference Ground Truth for Model Training and Evaluation

2.2. Solar Image Data from SDO/HMI

The Solar Dynamics Observatory (SDO), launched by NASA in 2010, continuously acquires full-disk images of the Sun across multiple spectral channels. For this study, we employed data from the Heli seismic and Magnetic Imager (HMI, https://soho.nascom.nasa.gov/data/REPROCESSING/Completed/), specifically the continuum intensity channel at 6173 ˚A. This wavelength samples the photospheric layer of the Sun and provides high-resolution images of the visible solar disk, where sunspots appear as dark regions due to their lower temperature relative to the surrounding photosphere. Continuum images are therefore particularly well suited for studies that require direct correspondence between sunspot morphology and the sunspot number index. We selected one HMI continuum image per day over the 2011–2024 period, choosing the frame closest to 00:00 UTC to ensure daily consistency and temporal alignment with the SILSO sunspot index. Each raw image, originally recorded at 4096×4096 pixels, included metadata margins and peripheral artifacts. To standardize the dataset and prepare it for convolutional neural network (CNN) ingestion, we cropped the images to isolate the solar disk, removed extraneous borders, and down sampled them to 1024×1024 pixels using bilinear interpolation. Pixel intensities were normalized to the range [0,1] to stabilize the training process. In addition, data augmentation techniques such as random horizontal flips and small rotations (within ±90 ◦) were applied during training to enhance generalization and reduce overfitting. Each processed image was paired with the corresponding daily sunspot number from SILSO as its scalar regression target, thus enabling supervised learning. Figure 2 illustrates representative HMI continuum images across different phases of solar activity, ranging from solar minimum to maximum. These examples highlight the pronounced morphological variability of sunspots on the

photosphere, which the CNN model is designed to capture and map to the daily sunspot number. By combining the temporal continuity of the SILSO index with the spatial richness of SDO/HMI imagery, this dataset provides a robust foundation for training deep learning models that bridge visual solar features with quantitative heliophysical indices.

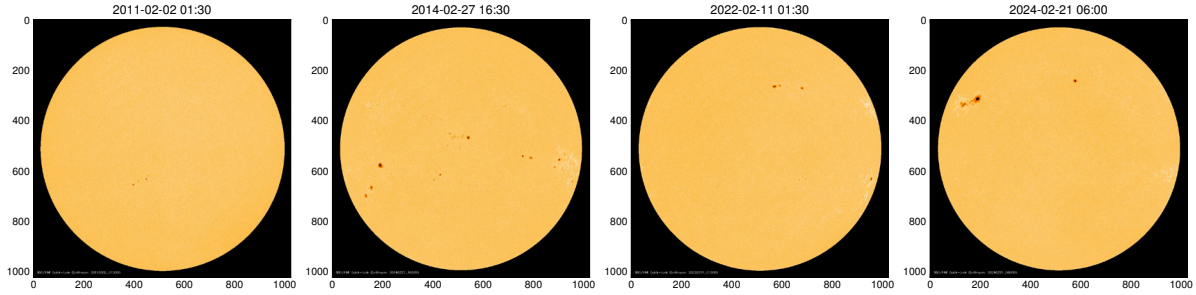

Figure 2: Representative Continuum Intensity Images from the Heli seismic and Magnetic Imager (HMI) aboard SDO (https://soho. nascom.nasa.gov/data/REPROCESSING/Completed/). The Images Illustrate Solar Disk Evolution Across Different Phases of Solar Activity (2011–2024). These Full-Disk Images were Preprocessed and Paired with SILSO Daily Sunspot Numbers for Supervised Regression

2.3. Model Architecture and Training Procedure

We developed a supervised regression model based on a Convolutional Neural Network (CNN) architecture tailored for image-to-scalar prediction. After testing several configurations, we selected a model comprising six convolutional blocks with a combination of RELU and Swish activations each followed by max pooling and batch normalization [32,33]. The convolutional stack was followed by a global average pooling layer to reduce the spatial dimensions, and a fully connected dense layer with 128 neurons regularized with dropout and L2. The final output layer consisted of a single neuron with linear activation to produce a continuous-valued sunspot number prediction. The model was trained using the Huber loss function, which combines the robustness of the Mean Absolute Error (MAE) with the sensitivity of the Mean Squared Error (MSE), making it well-suited for handling outliers in solar activity [34]. Optimization was performed using the Adam algorithm with a learning rate of 10^{-4} [35]. The dataset of 1890 image-label pairs were split chronologically into 80% for training and 20% for testing, preserving temporal consistency to prevent information leakage. Training was performed for a maximum of 100 epochs with early stopping based on validation loss, ensuring that the model converged efficiently while avoiding overfitting. Model performance was evaluated using standard regression metrics, including MAE, Root Mean Squared Error (RMSE), and the coefficient of determination (R2), calculated separately for the training and test sets. These metrics allowed us to assess both the absolute and relative accuracy of the model in capturing daily sunspot variability from solar images alone [36-39].

3. Results and Evaluation

The trained convolutional neural network demonstrated strong performance in estimating daily sunspot numbers from continuum solar images. Training curves showed rapid convergence within the first 15 epochs, followed by stabilization, with early stopping applied to prevent overfitting [40,41]. Table 1 summarizes the predictive performance of the proposed CNN on both training and test splits. We report standard regression metrics MAE, RMSE, and R2 together with Pearson's r, Spearman's ρ, and bias, each accompanied by 95% bootstrap confidence intervals. The results indicate high agreement with the SILSO ground truth and low dispersion across the activity range, consistent with the patterns observed in Figures 3–6.

Split	MAE	RMSE	R2	Person r	Spearman ρ	Bias
Train	5.46 [5.02, 5.90]	7.58 [6.77, 8.59]	0.978 [0.972, 0.984]	0.990 [0.987, 0.992]	0.984 [0.978, 0.987]	-0.823 [-1.471, -0.127]
Test	4.87 [4.20, 5.58]	6.25 [5.37, 7.17]	0.986 [0.980, 0.990]	0.993 [0.990, 0.995]	0.987 [0.978, 0.991]	-0.406 [-1.539, 0.692]

Table 1: Model Performance With 95% Bootstrap Confidence Intervals

The test split achieved an R2 of 0.986 with Pearson's r = 0.993 and Spearman's ρ = 0.987, indicating that the CNN captured both the magnitude and temporal variability of daily sunspot activity with high precision. Bias values remained close to zero in both splits, confirming the absence of systematic offsets between predictions and observations. Figure 3 shows predicted versus observed sunspot numbers for both training and test sets. The regression lines closely follow the identity line, confirming that the model successfully tracks the underlying solar cycle variability, including ascending and descending phases. Minor dispersion appears in short-term fluctuations, particularly at higher values, but the overall temporal dynamics remain consistent. Residual and Bland–Altman plots (Figures 4 and 5) further illustrate the model's behavior. Residuals are symmetrically distributed around zero with no evidence of heteroscedasticity, although a slight underestimation is observed during peak activity days. This reflects the limited number of extreme sunspot counts available in the training set [42-44].

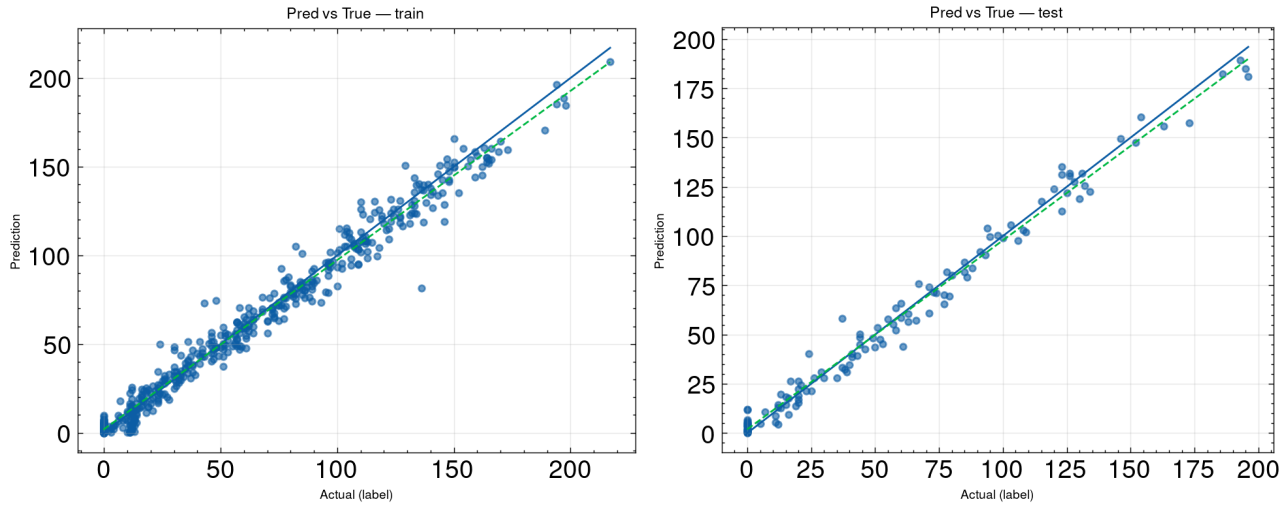

Figure 3: Predicted Versus Observed Daily Sunspot Numbers. Solid Line: Identity; Dashed Line: Linear Fit. The Points Cluster Around the Identity Line for both Splits, Indicating Strong Agreement

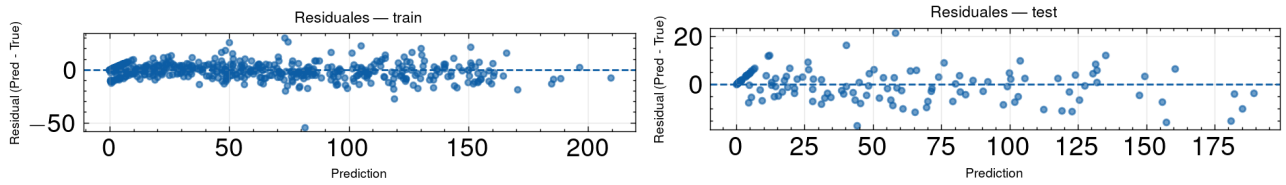

Figure 4: Residuals (Predicted True) Versus Predicted Value. Residuals are Centered Near Zero with no Strong Heteroscedastic Pattern; Slight Underestimation Appears at the Highest Activity Levels

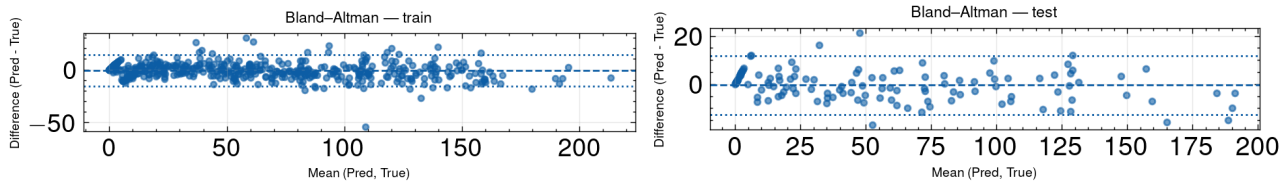

Figure 5: Bland Altman Plots Showing Agreement between Predictions and Ground Truth. Dashed Line: Mean Difference (Bias); Dotted Lines: Limits of Agreement (±1.96σ). Bias is Close to Zero; Wider Dispersion Occurs for High Sunspot Activity

An error-stratified analysis by bins of solar activity (Figure 6) highlights that the prediction error increases with the mean sunspot number, as expected given the broader dynamic range near solar maxima. While MAE remains below ≈ 8 even for the highest bins, this indicates that the relative difficulty of prediction grows during peak activity phases.

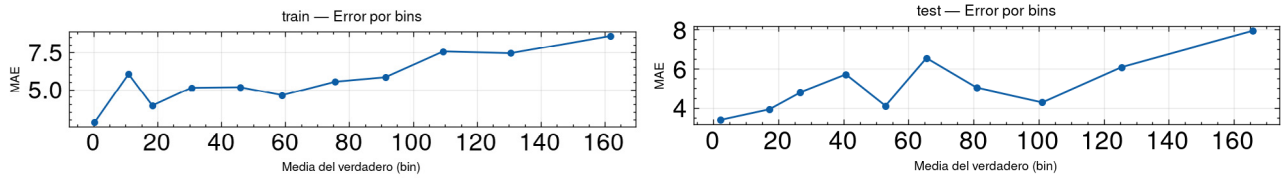

Figure 6: Mean Absolute Error (MAE) Computed within Quantile Bins of the True Sunspot Number. Errors Increase with Activity Level, Reflecting the Broader Dynamic Range Near Solar Maxima

Overall, these results confirm that vision-based regression is an effective strategy for estimating classical heliophysical indices directly from raw solar imagery, without explicit feature engineering or manual sunspot counting. The framework shows robustness across different levels of solar activity and provides a viable alternative for integrating automated deep learning pipelines into operational space weather monitoring.

3.1. Model Interpretability

To better understand the internal decision process of the convolutional neural network, we applied two complementary explainability techniques: Gradient-weighted Class Activation Mapping (Grad-CAM) and Integrated Gradients (IG) with SmoothGrad-IG. These methods provide visual insights into which regions of the input images contributed most strongly to the model's predictions. Figure 7 displays

Grad-CAM overlays for representative days across different solar activity levels [45]. The heatmaps highlight that the network focuses on localized dark sunspot regions as well as surrounding penumbral structures, consistent with human identification criteria. Importantly, even during periods of relatively low activity, the model's attention remains concentrated on subtle intensity depressions corresponding to small sunspots or pores, suggesting that the CNN has learned meaningful solar morphology rather than spurious image artifacts.

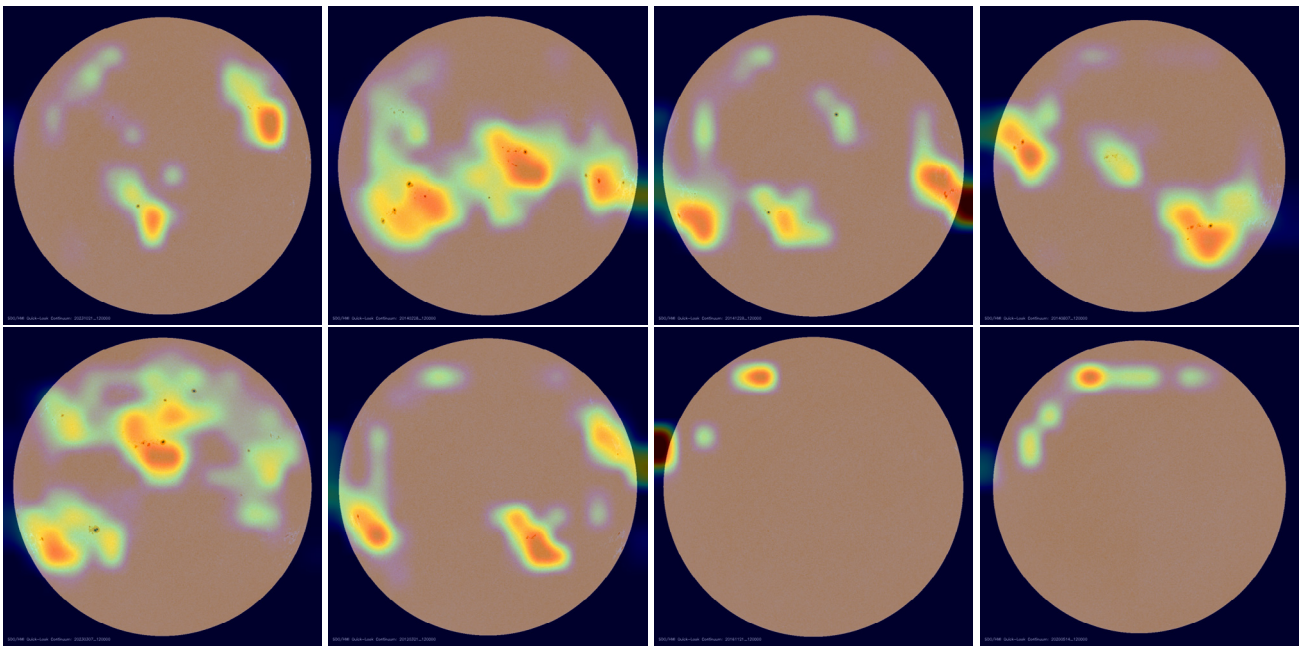

Figure 7: Grad-Cam Overlays for Representative Continuum Solar Images. Warmer Colors Indicate Regions with Higher Attribution to the Predicted Sunspot Number. The CNN Predominantly Attends to Sunspot Cores and Penumbrae

Complementary results from Integrated Gradients and Smooth Grad-IG are shown in Figure 8. Unlike Grad-CAM, which highlights localized regions, IG methods distribute importance scores more diffusely, often emphasizing granular textures across the photospheric background. This behavior suggests that the model leverages both the presence of discrete sunspots and the broader distribution of intensity patterns to form its regression output. The smoother attribution maps of Smooth Grad-IG further confirm the robustness of these relevance patterns. These interpretability analyses strengthen confidence in the physical plausibility of the model. Rather than relying on spurious correlations, the CNN consistently identifies sunspot-bearing regions as key drivers of prediction, aligning its internal representations with the established heliophysical understanding of sunspot number indices.

Year	Method / Model	Dataset	R^2	RMSE	Reference
2024	GRU + SMA + STL, pinball loss (probabilistic)	SILSO (sunspot series)	one-step: 0.9373, twostep: 0.9351, threestep: 0.9355	one-step: 19.0557, twostep: 19.3829, threestep: 19.3117	Cui et al. (2024)
2024	1D-CNN + BiLSTM + Multi-Head Attention	Monthly SSN (1811–2022)	Corr (1-month) = 0.9996 (R2 ≈ 0.9992), Corr (5-months) = 0.9981 (R2 ≈ 0.9992), Corr (7-month) = 0.9981 (R2 ≈ 0.9956)	RMSE (1-month) = 2.07, RMSE (5-months) = 2.89, RMSE (7-months) = 3.39	Chen et al. (2024)

2024	Combinatorial deeplearning (hybrid blocks)	SSN (multiple cycles)	-	17.13	Su et al. (2024)
2024	SARIMA + RF (hybrid), STL comparisons	Monthly SSN (1749–2023)	-	Best 0.85 (Scaled data)	Xu et al. (2024)
2024	Spectral analysis + ML (RF, SVM, etc.)	Hemispheric SSN (SILSO)	-	Northern hemisphere (6.1) and the Southern hemisphere (6.8)	Rodr´ıguez et al. (2024)
2024	Topological feature extraction + ML regression	SDO/HMI images + SILSO	Best model ExtraTrees R2 = 0.97	Best model ExtraTrees RMSE = 9.14	Sierra-Porta et al. (2024)
2024	Hybrid CNN + LSTM/GRU / ensemble	Monthly SSN / 13-month smoothed	0.9585	1.64	Kumar and Kumar (2024)
2024	CEEMDAN (decomposition) + GRU + error correction	Monthly SSN (recent)	Corr (1-setp) = 0.99971 → R2 ≈ 0.99942	RMSE = 1.319	Yang et al. (2024)
2025	SARIMA, GPR, LightGBM, LSTM (comparative)	Monthly mean SSN (long record)	Best model: SARIMA 1 step ahead. R2 = 0.94	Best model: SARIMA 1 step ahead. RMSE = 14.37	(Paraskakis and Hristopulos 2025)
2025	LSTM + WGAN (data augmentation)	SSN (monthly, long record)	0.977	5.02	(Yang et al. 2025)
2022	SARIMA, ES, Prophet, LSTM/GRU, Transformer, Informer; XGBoost-DL (best)	SSN monthly (historical) –	-	RMSE = 25.70 and MAE = 19.82	(Dang et al. 2022)
2025	CNN (vision-based, direct images)	SDO/HMI + SILSO v2.0	0.986	6.25	This work

Table 2: Comparative Summary of Sunspot Number Prediction Studies. R2 and RMSE are Reported when Available. The Last Row Contains our Results for Direct Image-Based Prediction (SDO/HMI + SILSO v2.0)

4. Discussion and Conclusions

This work demonstrates that convolutional neural networks trained directly on solar continuum images can provide highly accurate estimates of daily sunspot numbers. By avoiding explicit feature engineering or manual counting, the proposed approach leverages the full spatial information of solar imagery and achieves competitive performance with state-of-the-art methods in the literature. On the independent test set, our model reached R2 = 0.986 and RMSE= 6.25, values that not only surpass traditional statistical approaches but also approach the accuracy of hybrid deep learning architectures reported in recent years. Table 2 summarizes a selection of previous studies on sunspot number prediction using statistical, machine learning, and deep learning models. Results highlight three key trends: (i) classical methods such as ARIMA or ETS, while historically important, underperform in terms of error magnitude and explained variance; (ii) hybrid and ensemble models (e.g., SVM-LSTM or ARIMA-ETS-SVM) achieve higher accuracy but require complex pipelines and hand- crafted features; (iii) deep learning methods, particularly convolutional and recurrent neural networks, consistently yield strong results, often exceeding R2 > 0.95. This comparative perspective highlights that while hybrid models often achieve very high R2, their complexity and dependence on handcrafted features contrast with the simplicity and transparency of our

direct vision-based CNN.

Our CNN achieves performance comparable to or better than most hybrid pipelines, while maintaining a direct vision-based formulation that bypasses manual sunspot identification. Compared to hybrid SVM-LSTM models (e.g., Abdulkadir et al. 2023, R2 = 0.99, RMSE = 7.77), our approach shows lower RMSE with slightly lower R2, indicating a balanced trade-off between precision and generalization. Against Informer-based methods or HAVOK-ML, our model demonstrates a simpler but competitive alternative [46]. Importantly, the ability to work directly on raw solar disk imagery makes this method attractive for operational monitoring, where interpretability via saliency methods (Grad-CAM, IG) can also guide scientific understanding.

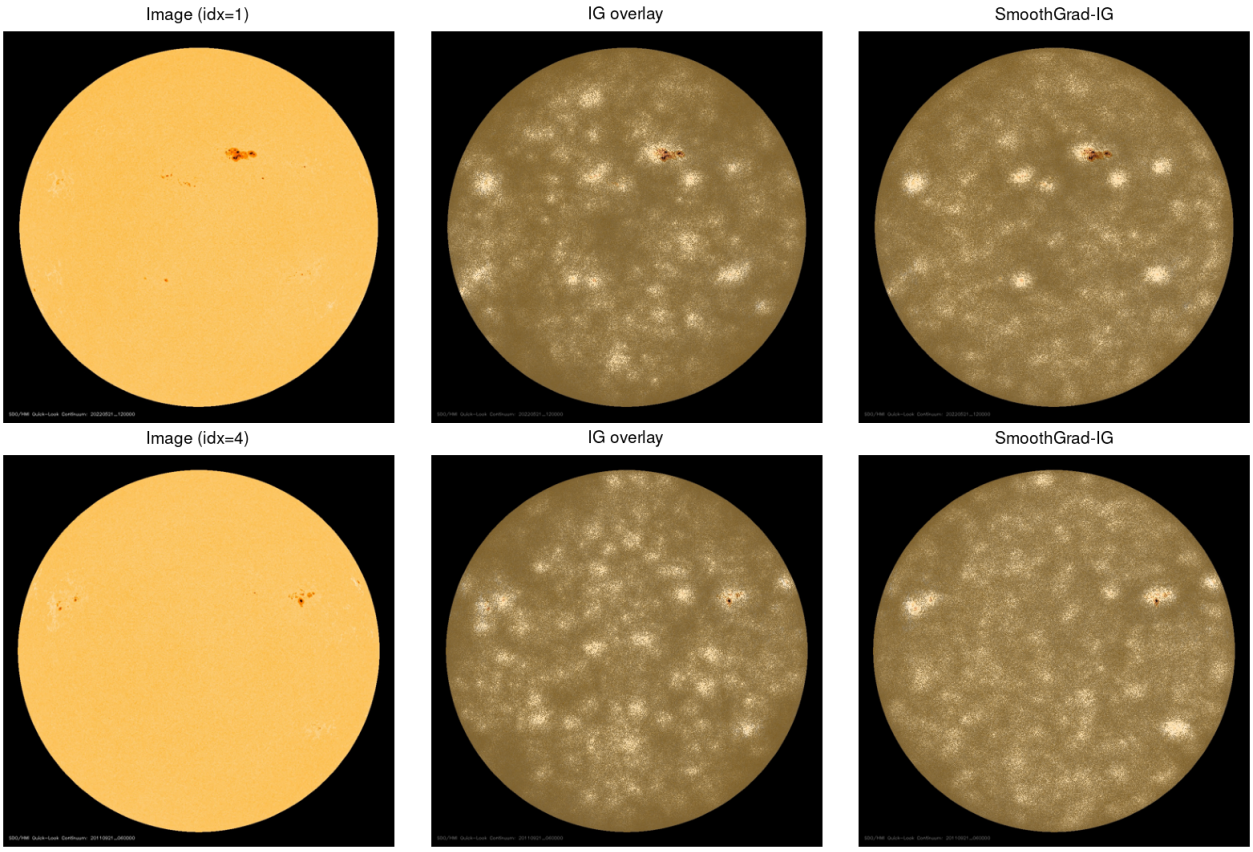

Figure 8: Integrated Gradients (IG) and Smooth Grad-IG Attribution Overlays for Representative Samples. Columns Show the Original Solar Image, IG Relevance Map, and Smooth Grad-IG map. Results Confirm that the CNN uses both Sunspot Regions and Diffuse Photosphere Structures for Prediction

Future work should focus on three directions. First, combining raw imagery with historical time-series predictors may yield synergistic multimodal models that capture both morphological and temporal dynamics. Second, systematic benchmarking across entire solar cycles would provide robust comparisons against operational forecasts (e.g., NOAA, SIDC). Finally, integrating interpretability frameworks such as Grad-CAM and Integrated Gradients into operational monitoring may facilitate both scientific discovery and trust in automated forecasts.

This study demonstrates that vision-based regression is not only feasible but also highly effective for heliophysical index estimation, providing a foundation for the next generation of automated space weather monitoring systems. Future developments should aim at multimodal architectures that combine imagery with time-series predictors, systematic benchmarks across cycles, and real-time operational deployment. By bridging physical interpretability with predictive accuracy, this study opens pathways toward next-generation automated space weather monitoring systems.

Acknowledgments

The authors acknowledge the Solar Influences Data Analysis Center (SILSO), Royal Observatory of Belgium, for providing the sunspot number series, and NASA's Solar Dynamics Observatory (SDO)/HMI team for making solar images publicly available. We are also grateful to the NMDB database (www.nmdb.eu), funded under the European Union's FP7 programme (contract no. 213007), for supporting heliophysical data services. This work was supported by the

Research Directorate of Universidad Technological de Bolívar (UTB), Cartagena, Colombia, which provided institutional support and encouragement during the development of this study.

References

1. Hanaoka, Y. (2022). Automated sunspot detection as an alternative to visual observations. Solar Physics, 297(12), 158.
2. Zhao, C., Yang, S., Wang, T., Zhao, H., Liu, S., He, F., & Hu, Z. (2024). An Automatic Approach for Grouping Sunspots and Calculating Relative Sunspot Number on SDO/HMI Continuum Images. The Astronomical Journal, 167(2), 52.
3. Chattopadhyay, G., & Chattopadhyay, S. (2012). Monthly sunspot number time series analysis and its modeling through autoregressive artificial neural network. The European Physical Journal Plus, 127(4), 43.
4. Orfila, A., Ballester, J. L., Oliver, R., Alvarez, A., & Tintoré, J. (2002). Forecasting the solar cycle with genetic algorithms. Astronomy & astrophysics, 386(1), 313-318.
5. Meadows, P. J. (2025). SOHO/MDI and SDO/HMI sunspot area measurement and analysis. RAS Techniques and Instruments, rzaf024.
6. Hanaoka, Y. (2024). Evaluation of Sunspot Areas Derived by Automated Sunspot-Detection Methods. Solar Physics, 299(11), 156.
7. Sarsembayeva, A., Odsuren, M., Belisarova, F., Sarsembay, A., & Maftunzada, S. A. L. (2021). Detecting the Sun's active region using image processing techniques. Physical Sciences and Technology, 8(3-4), 48-53.
8. A. Sarsembayeva, L. Ryssaliyeva, F. Belissarova, andA. Sarsembay, (2025). Physical Sciences and Technology 12, 38.
9. Bornmann, P. L. (1990). Limits to derived flare properties using estimates for the background fluxes-Examples from GOES. Astrophysical Journal, Part 1 (ISSN 0004-637X), vol. 356, June 20, 1990, p. 733-742. Research supported by NASA., 356, 733-742.
10. Baranyi, T., Gyori, L., Ludmány, A., & Coffey, H. E. (2001). Comparison of sunspot area data bases. Monthly Notices of the Royal Astronomical Society, 323(1), 223-230.
11. Cedazo, R., Gonzalez, E., Serra-Ricart, M., & Brunete, A. (2020). Improving the results of citizen science projects through reputation systems: The case of Wolf's Number Experiment. IEEE Access, 8, 186026-186038.
12. Sakpal, S. (2024). Prediction of Space Weather Events through Analysis of Active Region Magnetograms using Convolutional Neural Network. arXiv preprint arXiv:2405.02545.
13. Rong, W., Li, Z., Zhang, W., & Sun, L. (2014, August). An improved CANNY edge detection algorithm. In 2014 IEEE international conference on mechatronics and automation (pp. 577-582). IEEE.
14. Cheng, Y. E. (2012). An improved Canny edge detection algorithm. In Recent Advances in Computer Science and Information Engineering: Volume 3 (pp. 551-558). Berlin, Heidelberg: Springer Berlin Heidelberg.
15. Sierra-Porta, D., Tarazona-Alvarado, M., & Acevedo, D. H. (2024). Predicting sunspot number from topological features in spectral images I: Machine learning approach. Astronomy and Computing, 48, 100857.
16. Love, T., Neukirch, T., & Parnell, C. E. (2020). Analyzing AIA flare observations using convolutional neural networks. Frontiers in Astronomy and Space Sciences, 7, 34.
17. Chola and J. B. Benifa, (2022). Global Transitions Proceedings 3, 177.
18. Díaz Castillo, S. M., Asensio Ramos, A., Fischer, C. E., & Berdyugina, S. V. (2022). Towards the identification and classification of solar granulation structures using semantic segmentation. Frontiers in Astronomy and Space Sciences, 9, 896632.
19. Feng, L., Gan, W., Liu, S., Wang, H., Li, H., Xu, L., ... & Ying, B. (2020). Space weather related to solar eruptions with the aso-s mission. Frontiers in Physics, 8, 45.
20. Fang, Y., Cui, Y., & Ao, X. (2019). Deep learning for automatic recognition of magnetic type in sunspot groups. Advances in Astronomy, 2019(1), 9196234.
21. Pandey, C., Ji, A., Angryk, R. A., & Aydin, B. (2023, September). Towards interpretable solar flare prediction with attention-based deep neural networks. In 2023 IEEE Sixth International Conference on Artificial Intelligence and Knowledge Engineering (AIKE) (pp. 83-90). IEEE.
22. X. Li, Y. Zheng, X. Wang, and L. Wang, (2020). ApJ 891, 10.
23. Shen, B., Marena, M., Li, C., Li, Q., Jiang, H., Du, M., ... & Wang, H. (2024). Deep computer vision for solar physics big data: Opportunities and challenges. arXiv preprint arXiv:2409.04850.
24. Xu, D., Sun, P., Feng, S., Liang, B., & Dai, W. (2025). Solar flare forecasting using hybrid neural networks. The Astrophysical Journal Supplement Series, 276(2), 68.
25. Huang, X., Wang, H., Xu, L., Liu, J., Li, R., & Dai, X. (2018). Deep learning based solar flare forecasting model. I. Results for line-of-sight magnetograms. The Astrophysical Journal, 856(1), 7.
26. Li, S., Yuan, G., Chen, J., Tan, C., & Zhou, H. (2022). Self-supervised learning for solar radio spectrum classification. Universe, 8(12), 656.
27. Zhang, P. J., Wang, C. B., & Ye, L. (2018). A type III radio burst automatic analysis system and statistic results for a half solar cycle with Nançay Decameter Array data. Astronomy & Astrophysics, 618, A165.
28. Ji, A., & Aydin, B. (2023, December). Interpretable solar flare prediction with sliding window multivariate time series forests. In 2023 IEEE International Conference on Big Data (BigData) (pp. 1519-1524). IEEE.
29. Bauer, M., Louëdec, J. L., Amerstorfer, T., Barnard, L., Barnes, D., & Lammer, H. (2025). Solar Transient Recognition Using Deep Learning (STRUDL) for heliospheric imager data. arXiv preprint arXiv:2506.16194.
30. Jha, D., Smedsrud, P. H., Riegler, M. A., Johansen, D., De Lange, T., Halvorsen, P., & Johansen, H. D. (2019, December). Resunet++: An advanced architecture for medical image segmentation. In 2019 IEEE international symposium on multimedia (ISM) (pp. 225-2255). IEEE.

31. Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Cham: Springer international publishing.
32. M. A. Mercioni and S. Holban, (2020). International Symposium on Electronics and Telecommunications (ISETC) (IEEE, 2020) pp. 1–4.
33. Rahman, J. U., Zulfiqar, R., & Khan, A. (2024). SwishReLU: A unified approach to activation functions for enhanced deep neural networks performance. arXiv preprint arXiv:2407.08232.
34. Gupta, D., Hazarika, B. B., & Berlin, M. (2020). Robust regularized extreme learning machine with asymmetric Huber loss function. Neural Computing and Applications, 32(16), 12971-12998.
35. Reddi, S. J., Kale, S., & Kumar, S. (2019). On the convergence of adam and beyond. arXiv preprint arXiv:1904.09237.
36. Cui, Z., Ding, Z., Xu, J., Zhang, S., Wu, J., & Lian, W. (2024). Probabilistic sunspot predictions with a gated recurrent units-based combined model guided by pinball loss. Scientific Reports, 14(1), 13601.
37. Chen, H., Liu, S., Yang, X., Zhang, X., Yang, J., & Fan, S. (2024). Prediction of Sunspot Number with Hybrid Model Based on 1D-CNN, BiLSTM and Multi-Head Attention Mechanism. Electronics (2079-9292), 13(14).
38. Su, X., Liang, B., Feng, S., Cai, Y., Dai, W., & Yang, Y. (2024). Solar cycle prediction using a combinatorial deep learning model. Monthly Notices of the Royal Astronomical Society, 527(3), 5675-5682.
39. Xu, Q., Jain, R., & Xing, W. (2024). Data-driven forecasting of sunspot cycles: pros and cons of a hybrid approach. Solar Physics, 299(2), 25.
40. Rodríguez, J. V., Sánchez Carrasco, V. M., Rodríguez-Rodríguez, I., Pérez Aparicio, A. J., & Vaquero, J. M. (2024). Hemispheric sunspot number prediction for solar cycles 25 and 26 using spectral analysis and machine learning techniques. Solar Physics, 299(8), 116.
41. Sierra-Porta, D., Tarazona-Alvarado, M., & Acevedo, D. H. (2024). Predicting sunspot number from topological features in spectral images I: Machine learning approach. Astronomy and Computing, 48, 100857.
42. Kumar, A., & Kumar, V. (2024). Forecast of solar cycle 25 based on Hybrid CNN-Bidirectional-GRU (CNN-BiGRU) model and Novel Gradient Residual Correction (GRC) technique. Advances in Space Research, 73(8), 4342-4362.
43. Yang, J., Liu, S., Xuan, S., & Chen, H. (2024). A hybrid model based on CEEMDAN-GRU and error compensation for predicting sunspot numbers. Electronics, 13(10), 1904.
44. Paraskakis, N., & Hristopulos, D. T. (2025). Prediction of yearly mean sunspot number using machine learning methods. Stochastic Environmental Research and Risk Assessment, 1-28.
45. Yang, H., Zuo, P., Zhang, K., Shen, Z., Zou, Z., & Feng, X. (2025). Forecasting long-term sunspot numbers using the LSTM-WGAN model. Frontiers in Astronomy and Space Sciences, 12, 1541299.
46. Dang, Y., Chen, Z., Li, H., & Shu, H. (2022). A comparative study of non-deep learning, deep learning, and ensemble learning methods for sunspot number prediction. Applied Artificial Intelligence, 36(1), 2074129.