Predicting Coronal Hole Activity: Key to Mitigating Space Weather Impacts

Tarek Alsolame[1] tarek.alsolame@trincoll.edu,
BDS Aritra[1] bds.aritra@trincoll.edu,
Enock Niyonkuru[1] enock.niyonkuru@trincoll.edu,
Stephen Antogiovanni[1] stephen.antogiovanni@trincoll.edu, and
Chandranil Chakraborttii[1] nil.chakraborttii@trincoll.edu

Trinity College, 300 Summit Street, Hartford CT 06106, USA[1]

Abstract. Coronal holes significantly impact space weather, making understanding and predicting their activity crucial for mitigating effects on Earth's technological infrastructure. This paper presents a novel approach to forecasting short-term coronal hole activity using a multi-head CNN-LSTM model trained on Solar Dynamics Observatory (SDO) dataset tabular data and images. Our methodology involves preprocessing, training, and implementing a prediction system to predict the solar surface ratio covered by coronal holes. We compare our results with baseline models, including LSTM, GRU, and ARIMA-based models, demonstrating improved predictive performance. Sensitivity studies analyze parameter impacts on performance. Our work contributes to filling the gap in predictive modeling for coronal hole activity, providing valuable insights for space weather forecasting and preparedness. The RESNET-50 based LSTM-CNN model outperforms comparable models by 34% in MSE and 27% in MAE in predicting sunspot ratio, identifying optimal look back and predict ahead windows that further improve performance by up to 15%. Future work includes extending the model to longer-term predictions and incorporating additional features for enhanced accuracy.

Keywords: Coronal holes · Coronal hole activity prediction · Convolutional Neural Networks · Timeseries prediction.

1 Introduction

Coronal holes [7] are expansive regions within the Sun's corona characterized by reduced density and temperature. They hold significant scientific interest due to their open magnetic fields, which extend into space. This allows high-speed streams of solar wind to interact with Earth's magnetosphere, potentially triggering geomagnetic storms. Although such storms are commonly associated with solar minimums, they can occur at any time [6]. Coronal hole activity affects Earth in several ways, primarily through the emission of faster solar wind streams interacting with the magnetosphere, causing geomagnetic storms. These storms, categorized by National Oceanic and Atmospheric Administration (NOAA) [27],

can disrupt satellite operations, navigation systems, and radio communications, degrading high-frequency radio communications and increasing auroral activity visible even at lower latitudes [22]. Forecasters closely monitor coronal holes to predict their effects on space weather, which is vital for maintaining the reliability of space-dependent systems. Despite risks like disruption of communication systems and induction of electrical currents in power lines, the storms produced by coronal holes are typically less severe than those from coronal mass ejections (CMEs) [11]. Enhanced auroral activity from coronal hole activity can be visually stunning but also indicates disturbance in Earth's magnetosphere. While generally harmless on Earth's surface, intense auroral displays can pose risks to astronauts and space satellites due to increased radiation exposure. Despite their lesser severity, the long-lasting effects of coronal hole activity warrant advanced forecasting to prepare for potential impacts on technological systems and infrastructure. [6, 11, 27] .

Therefore, predicting coronal hole activity [21] is crucial for mitigating the effects of space weather on various technologies and services. These holes emit high-speed solar wind streams that, when interacting with Earth's magnetosphere, can cause geomagnetic storms with diverse impacts. These storms range from degrading high-frequency radio communications vital for aviation and maritime navigation to affecting satellite operations and GPS systems due to ionospheric changes. Furthermore, geomagnetic storms induced by coronal holes can disrupt power grids, leading to voltage instability or outages [31]. Predicting such events enables advanced warnings and protective measures, ensuring the resilience of technological systems and infrastructure. Astronomic forecasters monitor coronal holes closely to anticipate the strength and timing of solar wind streams, enhancing their ability to forecast geomagnetic conditions over the following days, crucial for space weather preparedness and the continuous operation of space-reliant technologies and services [33].

While the influence of coronal holes on space weather is widely acknowledged, there remains a gap in the literature concerning the prediction of their size and activity. While previous studies have predominantly focused on detecting increased coronal hole activity, accurately predicting such activity continously is crucial. The size and positioning of coronal holes directly impact the intensity and effects of solar wind streams on Earth [38]. However, limitations in previous research on coronal holes stem from observational uncertainties in determining their exact boundaries. Automated detection schemes like ASSA CH, CHARM, and CHIMERA [30] vary significantly in results due to diverse methods ranging from intensity-based thresholding to complex algorithmic classifications. Additionally, the dynamic nature of coronal hole boundaries, sites of constant interchange magnetic reconnection, makes them challenging to define uniquely. These challenges, coupled with the transient nature of coronal holes, make consistent monitoring complex. Addressing these limitations requires a cautious approach when using coronal hole boundary data for space weather predictions and modeling. Continued improvement in detection methods and a better understanding

of the physical dynamics at coronal hole boundaries are essential for advancing our knowledge in this area and improving space weather forecasting capabilities.

Recent research [14,26,33] on coronal hole activity has significantly advanced our understanding of coronal holes and their impact on space weather. SDO [29] was launched in 2010 to understand and record solar variations affecting Earth. It operates in a geosynchronous orbit, equipped with three scientific investigations: Atmospheric Imaging Assembly (AIA), Extreme Ultraviolet Variability Experiment (EVE), and Helioseismic and Magnetic Imager (HMI) [3]. SDO provides full-disk images of the Sun in various wavelengths, magnetic-field maps, and Doppler-velocity maps [34], offering unprecedented insights into solar dynamics and space weather.

In contrast to prior work that focuses on detecting abnormal coronal hole activity, the proposed novel approach focuses on predicting the size of coronal holes in the short term (1-7 days). We employ CHRONOS [13] to produce highlighted, masked versions of coronal holes, serving as ground truth. Then, we pre-process input images to compute the coronal hole coronal hole coverage ratio on the solar surface at 1-hour intervals (recorded as tabular data). The combined data, consisting of images and tabular data, is then fed into a multi-head CNN-LSTM model. The CNN-LSTM architecture combines the strengths of CNNs and LSTMs. CNNs excel at extracting features from input data, while LSTMs are adept at processing sequential information. This hybrid approach allows for effective sequence prediction by leveraging the spatial hierarchy of CNNs and the temporal learning capabilities of LSTMs to model intricate patterns in the data. The proposed approach, powered by the RESNET-50 CNN architecture, processes images and integrates both data types to predict future coronal hole activity with improved precision and accuracy. We describe our methodology in detail later in the paper. We present a comprehensive evaluation of various models for predicting sunspot ratio. Our results demonstrate the superior performance of the LSTM-CNN model, achieving the lowest Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) compared to other models. Sensitivity studies reveal that a look back period of 60 and a predict ahead window of 24 strike a balance between performance and computational overhead highlighting the effectiveness of the proposed method in capturing complex temporal patterns and its potential for applications in forecasting tasks.

2 Related Work

The Solar Dynamics Observatory (SDO) dataset [14] offers an unparalleled resource for solar physics research, providing high-resolution, multi-wavelength imagery of the Sun with exceptional temporal and spatial detail. The complexity and vastness of the SDO dataset make it ideal for image processing and machine learning techniques, which can uncover hidden patterns, improve predictions, and deepen our understanding of solar dynamics. Multiple works have leveraged this dataset to advance our understanding of solar phenomena and

develop innovative machine learning applications in heliophysics, including solar flare prediction using support vector machines [2], deep learning approaches for estimating coronal loop properties [10], flare precursor detection methods [5], and super-resolution imaging of magnetograms [14].

Research on coronal hole activity has traditionally emphasized detection methodologies, employing techniques ranging from manual inspection to automated computer vision methods to identify these regions on the sun [6, 7, 33]. Image processing techniques have been widely employed to detect coronal holes, characterized by cooler and less dense regions, using data from the observatories. Although successful in identifying their presence, these methods often fall short in providing predictive insights into their future behavior or evolution [14, 16]. Recent advancements [4,20] have shown promise in analyzing space weather data through deep learning approaches. Prior research utilized a introduced a deep learning method using the U-Net architecture to detect coronal holes [4, 17, 20], highlighting the potential of deep learning in this domain, however their work was limited to coronal hole segmentation, matching, and map classification.

Jarolim et al. [19] proposed the CHRONNOS model, a convolutional neural network (CNN) based technique that utilizes multi-channel data for precise coronal hole boundary detection [1], demonstrating the suitability of convolutional neural networks (CNNs) for real-time applications. Despite these improvements in detection accuracy, the focus has remained primarily on identification rather than prediction. CHRONNOS has been used in multiple works for various types of analysis on solar data such as solar irradiance variability on decadal to millennial timescales [9], revisit the solar forcing recommendations [12], etc. In this work, we utilize CHRONNOS to provide the ground truth information regarding coronal hole activity which we detail in a later section. Existing literature lacks robust predictive models for coronal holes, focusing mainly on detection rather than future behavior prediction. Our research introduces a novel approach using the LSTM architecture to forecast coronal hole activity over the short term period (1-7 days). This advancement addresses a critical gap in the field, offering valuable lead time for preparations against potential space weather events. Our work fills this void, promising enhanced predictive accuracy and positioning our research as a notable contribution to the field.

3 Proposed Approach

This section outlines our proposed approach, consisting of three key components: data preparation, model training, and the coronal hole prediction system. We describe the dataset, data preparation, model training process, and the end-to-end pipeline, including baseline models used for comparison.

3.1 Dataset

As mentioned earlier, we utilized the comprehensive dataset from NASA's SDO covering the years 2011 to 2018 [23]. This dataset includes a rich collection of solar images and related data, with the baseline timestamp-related data obtained

from prior work [19]. The dataset contained a CSV file that detailed temporal information for the period from 2011 to 2018, ensuring precise alignment with the corresponding image data. The SDO dataset comprises image matrices captured by the Atmospheric Imaging Assembly (AIA) and the Helioseismic and Magnetic Imager (HMI). Specifically, we included the AIA wavelengths of 94 Å, 131 Å, 171 Å, 193 Å, 211 Å, 304 Å, 335 Å, 1600 Å, and 1700 Å, as well as the HMI magnetic field components bx, by, and bz. For each hour, the file provides congruent paths to the respective image matrices, organized to facilitate seamless integration with the temporal data. The dataset was downloaded from the MiniSDO Dataset website [23] using the wget command. The data for each year provided as a .tar file that required extraction to access the individual image matrices. This extensive dataset (69 GB), combining high-resolution solar images across multiple wavelengths with precise temporal data, forms the foundation of our analysis, enabling a comprehensive study of solar phenomena, particularly focusing on the activity of coronal holes over time.

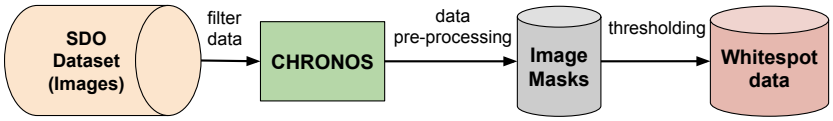

Fig. 1. Process for Computing sunspot Information

3.2 Data Preprocessing

Following the acquisition of the dataset, we employed CHRONNOS [19] to preprocess the raw images, utilizing its convolutional neural network to accurately delineate coronal hole boundaries. This pipeline provided access to hourly image data from the SDO dataset, allowing for the generation of binary mask images via thresholding (mean pixel values of the mask image) [24]. Subsequent sunspot computation [32] involved counting sunspot ratios and total pixels, yielding the sunspot ratio (ratio of white pixels in a solar image divided by total pixels). We then merged this tabular data with the original CSV file, creating a comprehensive dataframe that included relevant AIA wavelengths and SDO image paths. The sunspot ratios (tabular data) were then normalized using min-max scaling to a range of 0 to 1, ensuring that all features contribute equally to the model's learning process.

The image data was retrieved using paths extracted from our tabular CSV file, which also contained the sunspot ratio information. We focused primarily on the 171 Å wavelength image data due to its minimal missing values compared to other wavelengths used to produce CHRONNOS mask outputs. We selected the 171 Å wavelength for our analysis due to both its availability and its established relevance in observing coronal holes. While we acknowledge that other

wavelengths contribute valuable information about solar phenomena, the 171 Å channel offers several advantages. It captures emission from the upper transition region and quiet corona, providing a clear contrast between coronal holes (which appear darker) and the surrounding regions. Moreover, the 171 Å channel is less susceptible to contamination from bright active regions and flares, which can obscure coronal hole boundaries in other wavelengths. The relatively low number of missing values in the 171 Å data further supports its suitability for our study, ensuring a consistent and reliable dataset for training and evaluating our model. We analyzed missing values by examining their occurrence at an hourly cadence, revealing varying levels of missing data across different years. Due to the inherent complexity of interpolating image data, we proceeded with training our multi-head machine learning model using the available data, acknowledging the potential for future refinements through tabular data interpolation techniques. In addition to the raw image data and the sunspot ratios, no other features were explicitly extracted or engineered for this study. The focus was on leveraging the spatial information from the images and the temporal patterns in the sunspot ratios to predict coronal hole activity.

3.3 Model Architecture

The model employs a multi-head architecture, comprising both CNN and LSTM components, implemented using the PyTorch [18] and TensorFlow [28] library. The CNN-LSTM architecture [25] integrates the strengths of CNNs for feature extraction from input data with the sequential processing capabilities of LSTM networks, enabling effective sequence prediction. Originally designed for visual time series prediction tasks, such as generating textual descriptions from image sequences, this hybrid architecture leverages the spatial hierarchy of CNNs and the temporal learning capabilities of LSTMs to model complex patterns in data.

Our CNN-LSTM model employed RESNET-50 as a feature extractor, processing input images to extract spatial features like shapes, textures, and patterns through convolutional and pooling layers. The extracted features were then fed into an LSTM network, capturing temporal dynamics and patterns, such as changes in features over time. Our CNN-LSTM model, by combining RESNET-50's spatial feature extraction and LSTM's temporal modeling, effectively learns complex patterns in spatiotemporal data. We modify the RESNET-50 architecture to accommodate 12-channel input, corresponding to the 12 wavelengths of SDO images used per time step. The final fully connected layer of RESNET-50 is replaced with a linear layer that outputs 128 features per time step, followed by a ReLU activation function. A two-layer bidirectional LSTM with 500 hidden units processes the CNN embeddings sequentially, with a dropout layer with a probability of 0.25 applied after each LSTM layer to prevent overfitting. This ensures that the model generalizes well to unseen data. The bidirectional structure captures temporal dependencies in both forward and backward directions. A final dense layer with ReLU activation and linear output maps the LSTM's output to the prediction horizon, generating forecasts for multiple future time

steps. For instance, if predicting 10 future time steps, the dense layer's output dimension would be 10. The model architecture is described in Figure 2.

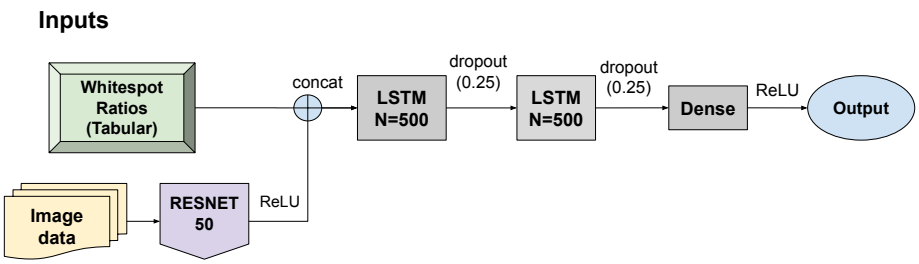

Fig. 2. Integrated Model Structure

3.4 Model Training

The inputs to our CNN-LSTM model include both tabular and image data (represented as matrices). Due to computational resource constraints, we employed a batched training approach, partitioning the dataset into yearly chunks and training the model on each year's data separately. This strategy enabled efficient resource utilization and allowed the model to adapt to distinct yearly patterns. Given the yearly variations in our data, limited computational resources, and temporal shifts in data distribution, this approach proved effective. Within each annual batch, we utilized a fixed batch size of 64, further dividing each year's data into smaller, manageable batches. This hybrid approach combined the benefits of annual batching, such as adaptability to yearly changes, with the control and stability offered by fixed batch sizes, enabling optimal training performance.

To mitigate the risk of overfitting, we employ several regularization techniques during training, including dropout layers with a probability of 0.25 after each LSTM layer and early stopping, monitoring the validation loss and terminating training if it fails to improve for a specified number of epochs. This helps to prevent the model from memorizing the training data and ensures that it generalizes well to unseen examples. The effectiveness of these regularization techniques is evident in the fact that our CNN-LSTM model did not trigger early stopping during training, suggesting that it has learned meaningful patterns in the data rather than overfitting to noise. We utilize the Huber loss function with delta=1.0 for training, striking a balance between Mean Squared Error (MSE) for small errors and Mean Absolute Error (MAE) for larger errors. To optimize the learning rate, we utilize a ReduceLR scheduler [15] with a minimum learning rate threshold, reducing the learning rate if the validation loss plateaus. We utilize the Adam optimizer [37] and early stopping with a patience of 10 epochs and monitoring of additional metrics (e.g., validation MAE) prevents overfitting.

The training data was reshaped to include a lookback and a predict-ahead parameter, which were adjusted during training for a comparative study. The lookback parameter indicated the recent history of number of hours of sunspot and image data used to train the LSTM cells used as context (12-84, in intervals of 12). The predict-ahead parameter, on the other hand, specified the number of hours the model was predicting into the future, with ranges of 24 to 168 in intervals of 24 (1-7 days). The model was trained on a high-performance computing infrastructure that included an Intel Xeon server (16 cores, 1.6 GHz, 96 GB DRAM) and an NVIDIA Tesla RTX 4090 GPU for accelerated training. For evaluation, we employed a temporal split of the dataset, allocating the first 80% of I/O accesses (ordered by timestamp) for training and the remaining 20% for testing. We evaluate the model's performance by comparing its predictions to the actual sunspot ratio, assessing its ability to capture the temporal dynamics of solar activity and generate accurate forecasts. Further analyses include comparing the model's RMSE to a baseline RMSE obtained by predicting the average target value, as well as examining the minimum and maximum slopes of the predictions to understand the model's sensitivity to changes in solar activity. This model architecture, combining the spatial feature extraction capabilities of CNNs with the temporal sequence learning capabilities of LSTMs, was designed to effectively capture and predict the dynamic patterns of coronal hole activity over time. By training the model in this manner, we aimed to achieve a high level of accuracy and reliability in our predictions, contributing valuable insights to the study of solar phenomena.

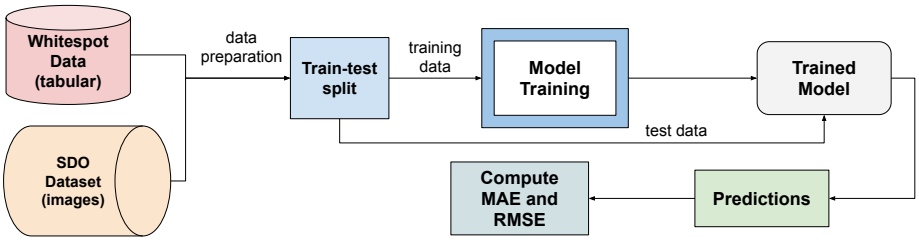

Fig. 3. Methodology

3.5 Coronal Hole Prediction System

Figure 3 illustrates the end-to-end coronal hole activity prediction system. Initially, raw image data from the SDO dataset was pre-processed to generate masked representations of solar images at different wavelengths using the CHRONNOS software. This process involved delineating the coronal hole boundaries and creating black-and-white mask images. We then computed the fraction of each image covered by coronal holes using a thresholding method [24] (Figure 1).

The resulting tabular data, which included the coronal hole coverage ratios, was combined with the image data collected at 171 Å. This combined dataset was used to data preparation for training the CNN-LSTM multi-head prediction model, which forecasts short-term coronal hole activity over a range of 1 to 7 days. We train the model offline using 80% of the data (training data), enabling it to make predictions on the remaining 20% (test data) within the system. To evaluate performance, we utilize Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Squared Error (MSE) metrics, which provide a comprehensive understanding of the model's accuracy and error distribution, allowing for comparison and benchmarking, and offering insights into strengths and weaknesses to guide future improvements. These metrics are well-suited for assessing the model's ability to predict coronal hole activity and identifying areas for refinement.

The model leverages RESNET-50 for spatial feature extraction and LSTM networks for capturing temporal dynamics, effectively integrating both spatial and temporal information to enhance predictive accuracy. To evaluate the impact of different lookback periods and forecasting horizons, we conducted a sensitivity and ablation study. This analysis involved varying the lookback periods (in 12 hour intervals) and forecasting horizons (in 24 hour intervals) to assess the robustness and reliability of our model. For evaluation, we compared the forecasted sun spot value with the ground truth in the tabular data obtained from CHRONNOS. Through this rigorous pre-processing and modeling approach, our system provides a powerful tool for predicting coronal hole activity, offering valuable insights for understanding solar dynamics and improving space weather forecasting.

3.6 Baselines

To establish robust baselines for evaluating our proposed CNN-LSTM model, we trained three distinct models: a standalone LSTM [35] model, which solely utilized the tabular data representing the fractional coverage of coronal holes on the sun's surface to forecast future coronal hole activity; a GRU [8] model; and an ARIMA-based model, renowned for its efficacy in modeling and predicting time-series data [36]. The LSTM baseline allowed us to assess the contribution of temporal dependencies in the tabular data, while the ARIMA baseline provided a comparison with a traditional statistical approach, widely employed in time-series forecasting. By comparing our proposed model's performance against these baselines, we can comprehensively evaluate its effectiveness in capturing complex coronal hole dynamics.

The selection of LSTM, GRU, and ARIMA as baseline models is grounded in their established efficacy and widespread use in time-series forecasting tasks. The LSTM and GRU models, representing recurrent neural network architectures, are particularly well-suited for capturing temporal dependencies in sequential data, making them relevant for modeling the evolution of coronal hole activity over time. The ARIMA model, a classical statistical approach, provides a robust benchmark rooted in traditional time-series analysis techniques. Furthermore,

these models have been extensively employed in solar physics research, offering a direct comparison with existing approaches in the field. While acknowledging the emergence of more recent deep learning architectures, the chosen baselines provide a solid foundation for evaluating the performance of our proposed model against established and widely recognized methods in both time-series forecasting and solar physics domains.

4 Results

In the following section, we present a comprehensive evaluation of the predictive performance of various models utilized in this study. The models under consideration include CNN-LSTM, LSTM, GRU, and ARIMA, each assessed based on their ability to predict sunspot ratio. We use Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) to evaluate model performance. RMSE is more interpretable than MSE, while MAE is less sensitive to outliers. MAE offers a robust measure, less sensitive to outliers. These complementary metrics provide a comprehensive understanding of model accuracy, robustness, and reliability.

Table 1. Results for different models (all years)

Model	**MSE**	**RMSE**	**MAE**	**Early Stopped**
CNN LSTM	4.22×10^{-4}	2.05×10^{-2}	2.18×10^{-2}	False
LSTM	1.21×10^{-3}	3.48×10^{-2}	2.60×10^{-2}	True
GRU	1.24×10^{-3}	3.52×10^{-2}	2.58×10^{-2}	True
ARIMA	1.53×10^{-3}	3.91×10^{-2}	2.85×10^{-2}	NA

4.1 Comparison with Baselines

The performance comparison of different models is summarized in Table 1. All models were trained using the initial 80% of the data from all years (2011-2018), while the remaining 20% was utilized for validation. The optimal values for the look back period and predict ahead window were determined using a grid search approach, where different combinations of these hyperparameters were evaluated based on their performance on the validation set.

The RESNET-50 based CNN-LSTM model demonstrates superior performance in predicting sunspot ratio, consistently outperforming other models across all evaluated metrics (MSE, RMSE, MAE). It achieves a 45% lower RMSE and a 42% lower MAE compared to the LSTM and GRU models. The improvement over the ARIMA model is even more substantial, with a 56% lower RMSE and a 58% lower MAE. This superior performance is particularly noticeable in years with high data variability, where the CNN-LSTM model maintains its

accuracy while other models struggle. For the experiments, we identified the optimal configuration by setting the look back period to 60 hours and the predict ahead period to 24 hours. This combination yielded the best performance, as demonstrated by our hyperparameter sensitivity study (shown later).

The CNN-LSTM model, leveraging both CNNs for spatial feature extraction and LSTMs for temporal dependencies, effectively handles the complex nature of sunspot ratio data. In contrast, LSTM and GRU models, while proficient in sequence modeling, show higher prediction errors, especially in years with significant fluctuations. Both also triggered early stopping, suggesting potential overfitting. The ARIMA model demonstrates the highest prediction errors, highlighting its limitations in capturing complex patterns. Due to computational constraints, with each epoch requiring over 3 hours of training, we limited the number of training epochs to 30. We also implemented early stopping, halting training if validation loss failed to improve by at least 10^{-5} for five consecutive epochs. Notably, the CNN-LSTM model did not trigger early stopping, indicating potential for further improvement with additional training or hyperparameter tuning. In contrast, the LSTM and GRU models may benefit from regularization or architectural modifications to enhance their performance. Our results underscore the significance of model selection and architecture in achieving accurate sunspot ratio predictions. The CNN-LSTM model stands out as the most promising candidate for further development and real-world applications, particularly in scenarios characterized by high data variability.

Table 2. Performance Comparison of Different Models Over the Years

Year	**RMSE**				**MAE**			
	Our Model	**LSTM**	**GRU**	**ARIMA**	**Our Model**	**LSTM**	**GRU**	**ARIMA**
2011	1.148e-01	2.084e-01	2.087e-01	2.145e-01	8.223e-02	1.568e-01	1.591e-01	1.605e-01
2012	2.403e-02	3.615e-02	2.873e-02	3.000e-02	1.387e-02	3.076e-02	2.237e-02	2.292e-02
2013	5.242e-02	1.009e-01	1.209e-01	1.036e-01	3.418e-02	7.113e-02	1.002e-01	7.209e-02
2014	1.027e-01	1.820e-01	1.806e-01	1.926e-01	6.983e-02	1.350e-01	1.341e-01	1.451e-01
2015	4.663e-02	8.005e-02	7.929e-02	8.147e-02	2.664e-02	5.940e-02	5.831e-02	6.005e-02
2016	4.753e-02	9.247e-02	9.288e-02	9.950e-02	3.248e-02	6.942e-02	6.946e-02	7.387e-02
2017	1.056e-01	2.015e-01	2.008e-01	2.130e-01	7.727e-02	1.545e-01	1.526e-01	1.650e-01
2018	8.986e-02	2.288e-01	2.245e-01	2.158e-01	6.585e-02	1.469e-01	1.478e-01	1.436e-01

4.2 Year-wise Performance Comparison

Table 2 presents a comprehensive comparison of our RESNET-50 CNN-LSTM model's performance, trained on 80% of the data for each year, against baseline models LSTM, GRU, and ARIMA. The results unequivocally demonstrate the RESNET-50 based model's superiority, consistently outperforming all baselines across all years in both RMSE and MAE metrics. This highlights the model's

exceptional predictive capability, showcasing its potential for accurate forecasting.

In particular, our model demonstrated a significant improvement over the baselines in years with high data variability, such as 2011, 2013, 2014, and 2017. In these years, the RMSE of the RESNET-50 based model was 42-53% lower than that of the LSTM and GRU models, and 49-53% lower than that of the ARIMA model. For instance, in 2011, RESNET-50 achieves an RMSE of 1.148×10^{-1} compared to the Base LSTM's 2.084×10^{-1}, Base GRU's 2.087×10^{-1}, and ARIMA's 2.145×10^{-1}. Similarly, in terms of MAE, RESNET-50 records 8.223×10^{-2}, while Base LSTM, Base GRU, and ARIMA show significantly higher errors of 1.568×10^{-1}, 1.591×10^{-1}, and 1.605×10^{-1}, respectively. These trends persist across subsequent years, highlighting the superior performance of the RESNET-50 LSTM-CNN model.

Table 3. Early Stopping Performance during Model Training

Year	**Early Stopped**		
	Our Model	**LSTM**	**GRU**
2011	False	True	True
2012	False	True	False
2013	False	False	True
2014	False	True	True
2015	False	True	True
2016	False	True	True
2017	False	True	True
2018	False	True	True

Furthermore, the early stopping criterion, implemented to prevent overfitting, was not triggered for the RESNET-50 based model in any of the years (Table 3), indicating that the model could potentially benefit from additional training. Its combination of convolutional and recurrent layers enabled effective capture of complex patterns, resulting in accurate predictions. In contrast, LSTM and GRU models frequently triggered early stopping, indicating higher susceptibility to overfitting and limited ability to learn complex patterns.

4.3 Sensitivity Studies

This section examines the impact of two key hyperparameters, look back and predict ahead, on the RESNET-50 based CNN-LSTM model's performance. Due to computational constraints, we focus on data from 2018, enabling efficient experimentation. We analyze the trade-offs between forecast accuracy, and training complexity, providing insights for optimizing model configuration.

4.4 Impact of Look Back

The look back period, determining the number of previous time steps considered by the model, significantly influences the model's performance. As can be seen in Table 4, increasing the look back period generally improves forecasting accuracy, with significant reductions in MSE, RMSE, and MAE scores. The most substantial improvement occurs between 48 and 60 hours, indicating a sweet spot for balancing longer-term dependencies and computational complexity. The key findings are:

- Increasing the look back period from 12 to 60 hours reduces MSE by 51.7%, RMSE by 31.0%, and MAE by 8.4%.
- The most significant improvement occurs between 48 and 60 hours, with a 19.1% reduction in MSE, a 26.6% reduction in RMSE, and a 4.4% reduction in MAE.
- Beyond 60 hours, the performance plateaus, indicating that incorporating more historical data does not lead to further improvements in forecasting accuracy.

These results suggest that a look back period of 60 hours strikes a balance between capturing longer-term dependencies and avoiding excessive computational complexity. Increasing the look back period from 12 to 60 hours reduces MSE by 51.7%, RMSE by 31.0%, and MAE by 8.4%. However, beyond 60 hours, the performance gains diminish, and the model's accuracy plateaus.

Table 4. Impact of Look Back on RESNET-50 LSTM-CNN Model Performance (2018)

Look back	**MSE**	**RMSE**	**MAE**
12	8.81×10^{-4}	2.97×10^{-2}	2.38×10^{-2}
24	7.61×10^{-4}	2.75×10^{-2}	2.17×10^{-2}
36	6.79×10^{-4}	2.60×10^{-2}	2.24×10^{-2}
48	5.22×10^{-4}	2.78×10^{-2}	2.28×10^{-2}
60	4.27×10^{-4}	2.09×10^{-2}	2.21×10^{-2}
72	5.21×10^{-4}	2.29×10^{-2}	2.51×10^{-2}
84	5.19×10^{-4}	2.28×10^{-2}	2.59×10^{-2}

4.5 Impact of Predict Ahead

The predict ahead parameter (PA), which determines the number of time steps into the future the model forecasts, significantly influences the performance of the RESNET-50 LSTM-CNN model. Table 5 presents the model's performance for various predict ahead values in 2018. As the predict-ahead value increases,

the model's predictive accuracy generally decreases, as reflected by the increasing MSE, RMSE, and MAE values. This degradation in performance is expected, as forecasting further into the future becomes inherently more challenging due to the accumulation of uncertainties and the potential for unforeseen events. Specifically, increasing the predict ahead from 24 to 48 hours results in a relatively small increase in error metrics (RMSE increases by 8.3%, MAE by 9.2%). However, further increasing the predict ahead to 72, 96, 120, and 144 hours leads to more substantial error increases (RMSE increases by 21.5%, 43.9%, 70.2%, and 107.3% respectively; MAE increases by 24.8%, 43.6%, 93.6%, and 117.4% respectively). These results indicate that the RESNET-50 LSTM-CNN model demonstrates good predictive capabilities for short-term forecasts (up to 48 hours ahead), but its accuracy diminishes considerably for longer-term predictions. This finding underscores the importance of selecting an appropriate predict ahead value based on the specific application and the desired balance between accuracy and forecast horizon.

Table 5. Impact of Predict Ahead on RESNET-50 LSTM-CNN Model Performance (2018)

Predict Ahead	**MSE**	**RMSE**	**MAE**
24	4.25×10^{-4}	2.05×10^{-2}	2.18×10^{-2}
48	4.95×10^{-4}	2.22×10^{-2}	2.38×10^{-2}
72	5.25×10^{-4}	2.29×10^{-2}	2.72×10^{-2}
96	9.97×10^{-4}	2.64×10^{-2}	3.13×10^{-2}
120	8.96×10^{-4}	2.99×10^{-2}	4.23×10^{-2}
144	1.18×10^{-3}	3.40×10^{-2}	4.73×10^{-2}

5 Conclusion and Future Work

In conclusion, this paper presents a pioneering approach to forecasting short-term coronal hole activity, leveraging a multi-head LSTM model that integrates tabular data and image matrices from the SDO dataset. Our comprehensive methodology yields promising results in predicting coronal hole activity over a 1-7 day period, outperforming baseline models, including an LSTM model using only tabular data and an ARIMA-based model. This study addresses a critical knowledge gap in the field by focusing on coronal hole activity prediction, which is essential for mitigating its impacts on space weather-dependent technological systems and services. Through rigorous sensitivity and ablation studies, we demonstrate the robustness of our model and identify areas for future improvement. Our model's promising results in predicting coronal hole dimensions are valuable for understanding space weather and its impact on Earth's technological infrastructure. The size and evolution of coronal holes are directly linked to

the properties of the solar wind streams they emit, which in turn can trigger geomagnetic storms and disrupt various technological systems.

Our work paves the way for future research directions, including extending the model to longer-term predictions, incorporating additional features, and exploring transfer learning applications. By advancing our predictive capabilities, we aim to contribute to developing more accurate and reliable space weather forecasting systems, ultimately enhancing the resilience of Earth's technological infrastructure against coronal hole activity impacts. Furthermore, we aim to establish a more explicit connection between our model's predictions and their physical implications, which could involve correlating predicted changes in coronal hole size with observed solar wind parameters, such as velocity and density, and investigating the relationship between these parameters and the occurrence and intensity of geomagnetic storms. By bridging this gap, we can enhance the practical utility of our model and contribute to the development of more effective space weather forecasting and mitigation strategies.

6 Acknowledgments

The authors wish to acknowledge the financial support provided by the NASA Space CT Grant Consortium under grant number Pro-Sum: P-2003. This funding enabled the completion of this research project and is greatly appreciated.

References

1. Abed, A.K., Qahwaji, R., Abed, A.: The automated prediction of solar flares from sdo images using deep learning. Advances in Space Research **67**(8), 2544–2557 (2021)
2. Bobra, M.G., Couvidat, S.: Solar flare prediction using sdo/hmi vector magnetic field data with a machine-learning algorithm. The Astrophysical Journal **798**(2), 135 (2015)
3. Boerner, P., Edwards, C., Lemen, J., Rausch, A., Schrijver, C., Shine, R., Shing, L., Stern, R., Tarbell, T., Title, A., et al.: Initial calibration of the atmospheric imaging assembly (aia) on the solar dynamics observatory (sdo). The Solar Dynamics Observatory pp. 41–66 (2012)
4. Chen, J., Deng, H., Li, S., Li, W., Chen, H., Chen, Y., Luo, B.: Ru-net: a residual u-net for automatic interplanetary coronal mass ejection detection. The Astrophysical Journal Supplement Series **259**(1), 8 (2022)
5. Chen, Y., Manchester, W.B., Hero, A.O., Toth, G., DuFumier, B., Zhou, T., Wang, X., Zhu, H., Sun, Z., Gombosi, T.I.: Identifying solar flare precursors using time series of sdo/hmi images and sharp parameters. Space weather **17**(10), 1404–1426 (2019)
6. Choi, Y., Moon, Y.J., Choi, S., Baek, J.H., Kim, S.S., Cho, K.S., Choe, G.: Statistical analysis of the relationships among coronal holes, corotating interaction regions, and geomagnetic storms. Solar Physics **254**, 311–323 (2009)
7. Cranmer, S.R.: Coronal holes. Living Reviews in Solar Physics **6**(1), 3 (2009)

8. Dey, R., Salem, F.M.: Gate-variants of gated recurrent unit (gru) neural networks. In: 2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS). pp. 1597–1600. IEEE (2017)
9. Egorova, T., Schmutz, W., Rozanov, E., Shapiro, A.I., Usoskin, I., Beer, J., Tagirov, R.V., Peter, T.: Revised historical solar irradiance forcing. Astronomy & Astrophysics **615**, A85 (2018)
10. Faber, J.T.: Characterizing Ultraviolet Bursts in a Solar Coronal Hole using Machine Learning Techniques. Master's thesis (2022)
11. Forbes, T.: A review on the genesis of coronal mass ejections. Journal of Geophysical Research: Space Physics **105**(A10), 23153–23165 (2000)
12. Funke, B., Dudok de Wit, T., Ermolli, I., Haberreiter, M., Kinnison, D., Marsh, D., Nesse, H., Seppälä, A., Sinnhuber, M., Usoskin, I.: Towards the definition of a solar forcing dataset for cmip7. Geoscientific Model Development Discussions **2023**, 1–14 (2023)
13. Funke, B., Dudok de Wit, T., Ermolli, I., Haberreiter, M., Kinnison, D., Marsh, D., Nesse, H., Seppälä, A., Sinnhuber, M., Usoskin, I.: Towards the definition of a solar forcing dataset for cmip7. Geoscientific Model Development **17**(3), 1217–1227 (2024)
14. Galvez, R., Fouhey, D.F., Jin, M., Szenicer, A., Muñoz-Jaramillo, A., Cheung, M.C., Wright, P.J., Bobra, M.G., Liu, Y., Mason, J., et al.: A machine-learning data set prepared from the nasa solar dynamics observatory mission. The Astrophysical Journal Supplement Series **242**(1), 7 (2019)
15. Hammoud, M., Sakr, M.F.: Locality-aware reduce task scheduling for mapreduce. In: 2011 IEEE Third International Conference on Cloud Computing Technology and Science. pp. 570–576. IEEE (2011)
16. Hurlburt, N., Cheung, M., Schrijver, C., Chang, L., Freeland, S., Green, S., Heck, C., Jaffey, A., Kobashi, A., Schiff, D., et al.: Heliophysics event knowledgebase for the solar dynamics observatory (sdo) and beyond. The solar dynamics observatory pp. 67–78 (2012)
17. Illarionov, E.A., Tlatov, A.G.: Segmentation of coronal holes in solar disc images with a convolutional neural network. Monthly Notices of the Royal Astronomical Society **481**(4), 5014–5021 (2018)
18. Imambi, S., Prakash, K.B., Kanagachidambaresan, G.: Pytorch. Programming with TensorFlow: solution for edge computing applications pp. 87–104 (2021)
19. Jarolim, R., Veronig, A., Hofmeister, S., Heinemann, S., Temmer, M., Podladchikova, T., Dissauer, K.: Multi-channel coronal hole detection with convolutional neural networks. Astronomy & Astrophysics **652**, A13 (2021)
20. Jatla, V., Pattichis, M.S., Arge, C.N.: Image processing methods for coronal hole segmentation, matching, and map classification. IEEE Transactions on Image Processing **29**, 1641–1653 (2019)
21. Joselyn, J., McIntosh, P.: Disappearing solar filaments: A useful predictor of geomagnetic activity. Journal of Geophysical Research: Space Physics **86**(A6), 4555–4564 (1981)
22. Korte, M., Stolze, S.: Variations in mid-latitude auroral activity during the h olocene. Archaeometry **58**(1), 159–176 (2016)
23. Kucuk, A., Banda, J.M., Angryk, R.A.: A large-scale solar dynamics observatory image dataset for computer vision applications. Scientific data **4**(1), 1–9 (2017)
24. Leung, C.K., Lam, F.: Maximum segmented image information thresholding. Graphical Models and Image Processing **60**(1), 57–76 (1998)
25. Livieris, I.E., Pintelas, E., Pintelas, P.: A cnn–lstm model for gold price time-series forecasting. Neural computing and applications **32**, 17351–17360 (2020)

26. McGregor, S., Dhuri, D., Berea, A., Muñoz-Jaramillo, A.: Flarenet: A deep learning framework for solar phenomena prediction. In: Neural Information Processing Systems (NIPS) 2017 workshop on Deep Learning for Physical Sciences (DLPS), Long Beach, CA, US (2017)
27. Neupert, W.M., Pizzo, V.: Solar coronal holes as sources of recurrent geomagnetic disturbances. Journal of Geophysical Research **79**(25), 3701–3709 (1974)
28. Pang, B., Nijkamp, E., Wu, Y.N.: Deep learning with tensorflow: A review. Journal of Educational and Behavioral Statistics **45**(2), 227–248 (2020)
29. Pesnell, W.D., Thompson, B.J., Chamberlin, P.: The solar dynamics observatory (SDO). Springer (2012)
30. Reiss, M.A., Muglach, K., Möstl, C., Arge, C.N., Bailey, R., Delouille, V., Garton, T.M., Hamada, A., Hofmeister, S., Illarionov, E., et al.: The observational uncertainty of coronal hole boundaries in automated detection schemes. The Astrophysical Journal **913**(1), 28 (2021)
31. Rust, D.M.: Coronal disturbances and their terrestrial effects: Tutorial lecture. Space Science Reviews **34**(1), 21–36 (1983)
32. Smith, K., Landes, P.E., Thollot, J., Myszkowski, K.: Apparent greyscale: A simple and fast conversion to perceptually accurate images and video. In: Computer graphics forum. vol. 27, pp. 193–200. Wiley Online Library (2008)
33. Vršnak, B., Temmer, M., Veronig, A.M.: Coronal holes and solar wind high-speed streams: I. forecasting the solar wind parameters. Solar Physics **240**(2), 315–330 (2007)
34. Yeo, K.L., Solanki, S.K., Krivova, N.A., Rempel, M., Anusha, L., Shapiro, A., Tagirov, R., Witzke, V.: The dimmest state of the sun. Geophysical Research Letters **47**(19), e2020GL090243 (2020)
35. Yu, Y., Si, X., Hu, C., Zhang, J.: A review of recurrent neural networks: Lstm cells and network architectures. Neural computation **31**(7), 1235–1270 (2019)
36. Zhang, G.P.: Time series forecasting using a hybrid arima and neural network model. Neurocomputing **50**, 159–175 (2003)
37. Zhang, Z.: Improved adam optimizer for deep neural networks. In: 2018 IEEE/ACM 26th international symposium on quality of service (IWQoS). pp. 1–2. Ieee (2018)
38. Zirker, J.B.: Coronal holes and high-speed wind streams. Reviews of Geophysics **15**(3), 257–269 (1977)