CJSJ

Improving Solar Flare Prediction using Deep Learning: Solar Flare Anticipation Algorithm (SOFAA)

Aaron Lee[1]

***Abstract*— High-energy solar superstorms can disrupt commercial, telecommunications, and energy infrastructure, posing a considerable risk to electronics in everyday life. Scientists have created deep-learning algorithms that can predict future solar superstorms, but these methods fail to incorporate numerical and visual data simultaneously. To increase the accuracy of deep-learning algorithms predicting solar flares, this paper proposes a deep-learning algorithm, SOFAA (Solar Flare Anticipation Algorithm), combining numerical telemetry and photos taken of the sun from the SDO (Solar Dynamics Observatory) and SOHO (Solar & Heliospheric Observatory) satellites. The SOFAA used data from the SOHO's total solar irradiance (TSI) measurements, measurements of energy from helium and hydrogen ions and electrons, and the SDO's solar images in the Fe IX ion spectrum. All data was taken in 2017, from January 1 to December 31. By utilizing GoogLeNet and multilayer perceptron models, the SOFAA achieved a loss value of 0.1255, demonstrating the ability to predict TSI values indicative of a solar flare a day in advance. In short, the incorporation of both numerical and visual data can give more accurate predictions of solar flare activity.**

I. Introduction

As the world becomes more reliant on electronic devices, the threat of solar superstorms grows increasingly relevant. High-energy solar superstorms—large ejections of charged particles and electromagnetic radiation [1]—exhibit a sporadic and hard-to-predict behavior that can threaten to shut down electrical grids; damage satellites, GPS systems, smartphones, solar cells, and telecommunication cables; disrupt avionics systems in airplanes; and disturb cellular, radio, and radar systems if proper measures are not taken to mitigate the effects [2, 3]. As such, scientists have made it an imperative to monitor the activity of the sun in order to predict the coming of a damaging superstorm. One of the ways scientists have analyzed data from tools like the Solar Dynamics Observatory (SDO) and the Solar & Heliospheric Observatory (SOHO) is by extracting patterns for predictions of storms through deep learning algorithms. Because these algorithms can detect hidden but noticeable patterns and relations between events, deep-learning programs are uniquely suited to solve the issue of solar flare and storm prediction [4]. However, previous models fail to make efficient use of all datasets available to them to make optimally accurate predictions.

Therefore, developing an accurate AI model for solar storm prediction with non-visual telemetry is vital to preventing the widespread destruction of the electrified global systems that support lives around the world. The Solar Flare Anticipation Algorithm, or SOFAA, was constructed in this study to respond to the rising awareness of the threat of high-energy solar storms on human transportation, communication, and commerce electronics. The aim of this paper is to improve prediction of solar storms by combining visual and non-visual data elements from the sun.

II. Materials and Methods

To create the training data set, this study used data from the SDO's AIA channel 171, recording the electromagnetic radiation from the ion Fe IX to provide visual data for the deep learning program's image processing [5]. The imaged wavelength provided simple gold-scale contrast for the program to derive patterns, specifically that of how the appearance of the sun prior to solar events changes from baseline solar appearance. The numerical data from the SOHO's COSTEP/EPHIN instruments were used to analyze how the energy of hydrogen ions emitted by the sun shifts as a solar event draws near and occurs, while the same from the SOHO's VIRGO was used as an indicator of solar flare presence, as average TSI peaks during solar flare events [6]. All data was derived from 1 January 2017 to 31 December 2017 because an X8.2 level solar flare occurred in 2017, serving as a good training dataset for the appearance of a large solar flare [7]. VIRGO instrument data was taken hourly; COSTEP/EPHIN instrument data was taken every five minutes; and AIA instrument data was taken about every three minutes. All data points from the COSTEP/EPHIN instrument and AIA instrument were averaged and associated with an hour of a day in accordance with the VIRGO instrument data to sync the data up with each other.

The perceptron, proposed by Frank Rosenblatt in 1957 and serving as the basic building block of deep-learning algorithms, is defined by a single equation, $Wx + b$, where W is weight and b is bias. The perceptron randomly sets a weight and bias value, inputs a training data input value for x, calculates the output, finds the difference between the predicted output and actual training data output, and performs those steps for each input value until the differences in all predicted outputs and actual outputs can be calculated into a loss value. The algorithm then resets the weight and bias values in a direction that reduces the loss value and repeats the previous steps again, effectively reducing the loss value to replicate a pattern in the data set. However, the line defined by the perceptron equation cannot predict all patterns as Rosenblatt thought, so a more complex structure of many perceptrons was needed.

[1]Aaron Lee is with the Meadow School, Las Vegas, NV 89128 USA (e-mail: aaronjlee0422@gmail.com).

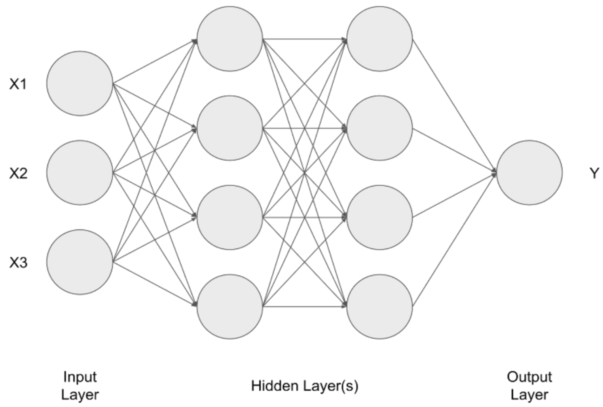

Figure 1. Multi-layer perceptron layer, input, and output structure

To resolve the shortcomings of the single-layer perceptron, the multilayer perceptron, or MLP, was invented. An MLP is composed of many perceptrons interconnected in webs with each other and one or more inputs and outputs. An MLP is organized into three "layers": an input layer that processes the raw input data, hidden layers that further process the input layer results, and an output layer that transforms the values from the hidden layer into meaningful output values. Due to the scalable nature of MLPs, they are able to derive complex patterns from any numerical data set.

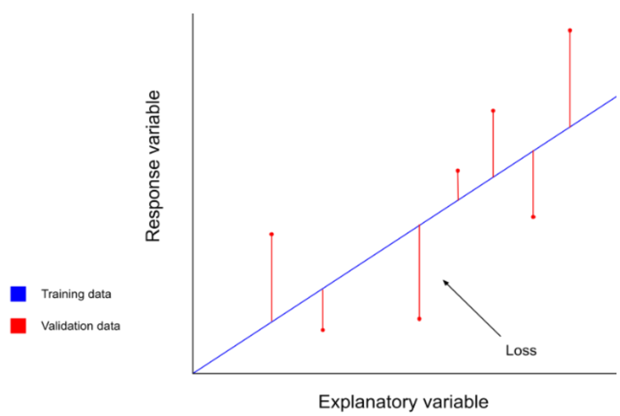

Figure 2. Deep-learning algorithm loss calculation method; blue line is training data set, red dots are predictions by the program, and red lines are the difference between the predicted data and the actual training data, also known as loss

Loss functions are mathematical functions that allow algorithms to calculate loss values measuring the deviation that the program predictions have from the training data presented. There are several types of loss functions, two of which are mean absolute error, MAE, and mean squared error, MSE. The following equation models MAE, given n is the sample size, y_i is the prediction value, and x_i is the actual value:

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}$$

The following equation models MSE, given the same variables above:

$$MSE = \frac{\sum_{i=1}^{n} (y_i - x_i)^2}{n}$$

The MAE loss function tends to adapt to large outliers less and gives greater corrective weight to small errors, while the MSE loss function amplifies and thus corrects to large outliers more due to squaring results. Regardless of the method, loss functions turn the comparisons between the outputs of the algorithm and those of the training data set into a score, the loss value, which the algorithm attempts to lower by adjusting the weight and bias values.

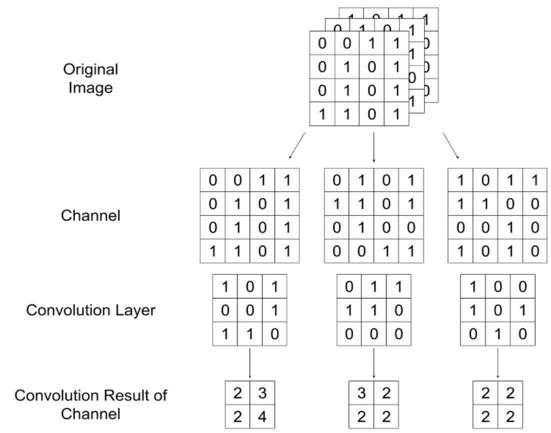

Figure 3. Convolutional layer processes; images are broken apart into three color channels, the convolution layer values are stacked and multiplied to every value underneath it at every possible position, and the resultant multiplied values are added up to create one value in the result matrix.

A convolutional neural network, or CNN, is an algorithm that replicates the way in which the human brain processes visual information. Like the actual brain, CNNs are used by computers to process images in a way understandable to itself. CNNs are made up of three parts: a convolutional layer, which breaks the image into RGB channels and runs the pixel values through filter matrices, a pooling layer, which removes outlier values skewing patterns from the final convolutional layer results, and a fully connected layer, an MLP that finds patterns in the numerical values of the final result matrix. It is through this algorithm that images can be processed by computers.

GoogLeNet is Google's rendition of a CNN that resolves previous issues with CNNs overfitting, where the algorithm finds patterns too specific to the training data set. GoogLeNet averts overfitting by diversifying the module architectures that make up the CNN's convolutional layer, removing repetitive and pointless actions in the CNN to ensure faster pattern retrieval. GoogLeNet is implemented into the Tensorflow-Keras AI program [8].

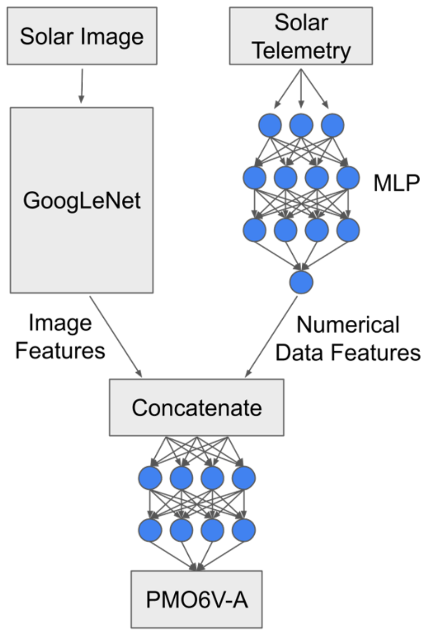

Figure 4. SOFAA deep-learning algorithm architecture

SOFAA is a combination of the GoogLeNet CNN and the individual MLP architectures. The SOFAA runs the SDO satellite's AIA solar images through the GoogLeNet CNN and the SOHO satellite's COSTEP proton energy level measurements through an MLP to get so-called "features," or patterns, from the input data. These features are concatenated, or combined, into a single feature set, which is then run through another MLP to get the final predicted TSI measurements to be compared with those of the SOHO satellite's VIRGO instrument. Three variables were changed to see their effects on the resultant loss value: loss function used, sample size, and dense layer arrangements and types. The loss values used to evaluate the effectiveness of variable changes were run through another MAE function to create comparable values. The training data TSI values were pushed forward one day to test for the SOFAA's ability to predict TSI values one day into the future.

III. Results

The loss values for all trials of the SOFAA model were less than 1, indicating high correlation between solar imaging patterns, charged proton intensity telemetry values, and TSI values from the sun. Differing dense layer sizes also did not affect the amount of loss experienced with the model. For the most part, an increase in sample size resulted in lower loss values, although by an arguably negligible amount. Loss remained consistent after slight variation in loss in the first epoch. This may indicate high correlation in the values of the three datasets, as finding patterns was of such ease that the ideal weight and bias values resulted after just one epoch. For sake of brevity, the number of epochs was reduced to 7.

TABLE I. SOFAA model experimentation results (7 epochs)

Trial #	Data point sample size	Dense layers	MAE/MSE	Loss Value
1	2000	100-300-100	MAE	0.3329
2	2000	100-300-100	MSE	0.1283
3	2000	100-300-900-500-100-100	MAE	0.3329
4	2000	100-300-900-500-100-100	MSE	0.1283
5	6000	100-300-100	MSE	0.1274
6	10000	100-300-100	MSE	0.1255

After verifying the ability of the program to find correlations between data taken at the same time, experiments displacing the TSI data by one day forward were performed. With 7 epochs, a 10000 data point sample size, 100-300-100 format dense layers, and an MSE loss calculation method, the SOFAA model was able to find synergy between solar imaging, proton intensity telemetry measurements, and TSI measurements a day into the future, as low loss indicates the ability of the program to find a pattern or correlation in the data input and output. The lowest loss value achieved was 0.1257.

Figure 5. SOFAA model experimentation results chart

IV. Conclusion

In this study, the SOFAA model was created and refined with the goal of combining visual and numerical solar data to increase the accuracy of solar flare predictions. With the advancement of technology in the modern era, many more opportunities to ensure the security of human infrastructure have presented themselves. The SOFAA's results show promise in its ability to predict the value of TSI one day into the future when given solar imaging and proton energy intensity value data, predicting within a reasonable margin of error and loss of 0.1257. With the predicted TSI measurement, the determination of whether a solar flare will occur or not and of what intensity can be derived with a large amount of confidence, allowing for advance notice of solar flares.

V. Discussion

The SOFAA helps illuminate visual and numerical data concatenation as an important path to improving solar flare prediction moving forward. Most solar flare prediction model proposals still restrict themselves to the analysis of visual data and qualitative feature sets. Through the correlation found between the concatenated data set and the future TSI value of the sun here, the creation of a prediction model for solar flares utilizing concatenated data sets is not far away.

However, despite the successes of the SOFAA model, there remains ample room to expand and further push the model to serve greater and more demanding purposes. In its current state, the prediction values for TSI from the sun must be analyzed manually to check for the probability of a solar flare or storm occurring. In most cases, this analysis would still require additional processing time. Another set of algorithms to further derive patterns in the TSI measurements to determine solar superstorm probability would solve this issue efficiently and quickly. Although the SOFAA model has low loss values in its predictions, there may still be concern about the verity of such results. However, due to the unknowable nature of the mechanisms in deep-learning algorithms, deriving an easy way to further analyze and interpret the resulting data is difficult. One way to increase certainty in the results besides directly analyzing the data is to cross-examine the results of the SOFAA model with other models of a similar nature, like that of the Deep Flare Net (DeFN) model, to check for similar results and rule out discrepancies manually as needed. Another way to improve the SOFAA model's verity would be to add more datasets into the pool of solar telemetry data used by SOFAA. By adding more solar telemetry data in the model, SOFAA would be able to further ensure that the patterns it does detect are noticeable and common across all analyzable elements of the sun. One more method to improve accuracy of SOFAA model result data would be to utilize the true skill statistics (TSS), a metric of discrimination performance, used to evaluate the DeFN model, as the TSS skill score considers many more factors, including loss, to help give an operational "skill score" of the TSS reflective of actual prediction performance in practical situations [9]. Overall, the SOFAA model demonstrates the great potential for solar flare prediction software improvement, which lies in the concatenation of different data set types by reinforcing trends present in both data sets and diminishing those only present in singular data sets.

VI. References

[1] J. Hu, V. S. Airapetian, G. Li, G. Zank, and M. Jin, "Extreme energetic particle events by superflare-associated CMEs from solar-like stars," *Science Advances*, vol. 8, no. 12, Mar. 2022.

[2] P. Cannon, M. Angling, L. Barclay, C. Curry, C. Dyer, R. Edwards, G. Greene, M. Hapgood, R. Horne, D. Jackson, C. Mitchell, J. Owen, A. Richards, C. Rogers, K. Ryden, S. Saunders, M. Sweeting, R. Tanner, A. Thomson, and C. Underwood, "Extreme space weather: impacts on engineered systems and infrastructure," *Royal Academy of Engineering*, Feb. 2013.

[3] Atmospheric and Environmental Research, "Solar Storm Risk to the North American Electric Grid," *Lloyd's*, 2013.

[4] E. Larsen, "Machine Learning Model Survey with the dataset for solar flare prediction," *PeopleTec, Inc.*, 2021.

[5] M. Lockheed. (n.d.). "Instrument. Atmospheric Imaging Assembly," https://aia.lmsal.com/public/instrument.htm

[6] M. Kretzschmar, T. D. d. Wit, W. Schmutz, S. Mekaoui, J. F. Hochedez, and S. Dewitte. "The effect of flares on total solar irradiance," *Nature Physics*, vol. 6, pp. 690-692, Aug. 2010.

[7] J. Zhao, W. Liu, and J. C. Vial, "White-light Continuum Observation of the Off-limb Loops of the SOL2017-09- 10 X8.2 Flare: Temporal and Spatial Variations," *The Astrophysical Journal Letters*, vol. 921, no. 2, Nov. 2021.

[8] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going Deeper with Convolutions," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1-9, 2015.

[9] N. Nishizuka, Y. Kubo, K. Sugiura, M. Den, and M. Ishii, "Deep Flare Net (DeFN) Model for Solar Flare Prediction," *The Astrophysical Journal*, vol. 858, no. 113, May 2013.