ISSN 2616-5775 Vol. 6, Issue 8: 158-164, DOI: 10.25236/AJCIS.2023.060819

Prediction Based on Convolutional Neural Networks and Vision Transformer for GOES-XRS Solar Flare Time Series

Mingjie Wei

Shanghai Vanke School Pudong, Shanghai, China

Abstract: *Solar flare is a type of solar activity that occurs at active regions at the surface of the sun. The emission of solar flares has numerous consequences, including the dis- turbances of magnetic fields, disruptions from energetic particles, and geomagnetic explosions. All those consequences have numerous impacts on human civilization, including the degradation of communication systems, power grids, space navigation, and even natural disasters. Thus, those minor or catastrophic consequences are al- ways threatening to the normal operation of society and decision-makers of those systems always seek a precise and accurate prediction of hazardous solar flares. This paper aims to develop a forecast model that can accurately decide whether solar flares would happen in the future. The data is extracted from the NOAA (National Oceanic and Atmospheric Administration) GOES-16 X-Ray Sensor that monitors solar activity by measuring the flux intensity of X-Ray. The original data is in the form of time series. Markov Transition Field is applied to the time series data, transforming the data into the form of 3-dimensional images. Therefore, the data undergone pre-processing could be applied to computer vision models. The aim of these models is to accurately recognize the Markov Transition Field images that symbolize there would be solar flare emission one hour later through a binary classification. Deep learning architects are the major components to accomplish this forecast task. Convolutional Neural Network (CNN) is a common approach in doing clas- sification tasks, which is also frequently used in recent studies that aim to predict flare emission through X-Ray images of the active regions. There are several classic CNN undergone training and testing, including LeNet-5, AlexNet, VGGNet 16 and 19, and ResNet-18, that utilizes the residue block structure. These CNN architects provide fascinating reliability and accuracy in this prediction task of solar flares, with multiple structures providing accuracy greater than 80%. Furthermore, Vision Transformer, a deep learning architect also used in classification based on trans- former structure is applied to the flare task. It is comprised of the core structure of multiple-head self-attention, residue blocks, layer normalization, and multilayer perceptron. Vision Transformer has shown outstanding accuracy (89.89%) while making predictions of solar flare emissions.*

Keywords: *Solar flare, NOAA XRS, GOES-16 satellite, Markov Transition field*

1. Introduction

The sun, the only star in our solar system, has a mass of 1.9891×1030 kg. It is constituted of hydrogen, which is the majority, and other elements such as helium. The sun is interwound with plasma and magnetic fields, forming an ideal spherical celestial body. Hydrogen in the sun undergoes nuclear fusion, releasing heat and light to space, and forms helium.

Solar activities occur on the surface of the sun frequently and ceaselessly. The solar wind is a solar activity comprising subatomic particles including protons and electrons moving at high speed. It ejects particle flows into space. There are also several types of solar activities triggered by solar magnetic fields. Sunspots appear as dark spots on the solar photosphere are the aggregation of magnetic fields. They have areas from 500 to 1000 kilometers. They usually have a temperature of around 4000 Celsius degrees, which is lower than the average temperature of the surface of the sun, making them have less luminosity than the surrounding. The intense aggregation of magnetic fields at the sunspots would facilitate the formation of solar flares.

Solar flares are massive ejections of energy from the solar surface, caused by magnetic reconnection. They are frequently observed above sunspot groups. As the assembled magnetic fields are more complicated, it the easier could reserve magnetic energy. When magnetic energy at a certain area, such as a sunspot group, is in excess, energy would be released to space. The ejection of energy takes form in radiations of a full wave band and it is accompanied by high-energy particle flows (from 103eV to 109eV) and impulse waves. The majority of radiations are in short wavelengths, such as X-ray and ultraviolet radiations[1-4].

Based on different observation approaches, there are mainly two methods of classifi- cation of solar flares. The standard classification method, which is used on ground-based optical images, classifies solar flares with a numerical value from 0 to 4 and a letter (F, N, or B) with respect to their optical intensity. Optical observations usually detect a sudden amplification of the optical waveband, usually having wavelengths between 3900 to 7000 angstroms. The classification defined by GOES (Geostationary Operational Envi- ronmental Satellite) observations has divided the intensity of Solar flares into 5 increasing levels, A, B, C, M, and X. This classification is based on the peak flux of X-ray, which a wavelength between 0.01 to 100 angstroms is detected by the XRS (X-Ray Sensor). Solar flare emissions, especially with an intensity greater than M-level, are usually accompanied by strong proton flows. Since Earth's ionosphere is more sensitive to the variation of the flux of X-ray radiation, it's the prevalence that using X-ray classification standards to define the intensity of solar flares. Therefore, this research focuses on determining solar flares based on the X-ray flares scale. Information in this section has consulted papers regarding GOES satellites such as [5].

Solar flares emissions follow a period of 9 to 14 years (an average of 11 years), forming solar cycles. Contemporarily, it is the 25th observed solar cycle, the observation started in 1755, and the 25th cycle starts in 2019.

Generally, normal solar flares between levels A to B have minor influences on either advanced technologies and infrastructures. However, more intensified solar flares (X and M flares) could bring significant impacts on human civilization. After reaching Earth's ionosphere (a highly ionized layer), it reflects radio signals and prompts the communica- tion system), radiation from solar flares would increase the ionization of the ionosphere. Hence, radio signals will be more likely absorbed (and not reflected), resulting in a reduc- tion in the reliability of radio communications. If the flux of X-ray reaches some extent, it also influences receptors of all communication devices, including civilian and military. A strong and rapid disturbance in the magnetic field, caused by the solar flare, could induce an intense current within the components of the devices. If the device lacks protection in such circumstances, those induced currents would increase the voltage and can result in irreversible damage. A similar mechanism damages power grids (if protections could not negate flares' effects). For long-range high-voltage transmissions, the sudden intensifica- tion of electromotive force might overwhelm the load of power grids that is usually at full capacity in particular regions. Overload would cause extreme and irreversible degradation to the power supply.

For satellites, the high-energy particles from the solar flare could strip electrons out of the satellite's components. Once the electrons are delocalized, they would cause a short circuit of the electronic components. Moreover, the geomagnetic explosion (dra- matic Earth magnetic field disturbance) caused by solar activities would influence the torque of operating satellites, resulting in the alteration of their position or direction.

The disturbances on the ionosphere and upper atmosphere also increase the density of the upper atmosphere, which increases the resistance of the movement of satellites in their orbits. The effects on satellites would lead to the inaccuracy of GPS (Global Po- sition System), resulting in uncertainty of even dozens of kilometers. Those impacts on communications especially constitute a danger for civilian aviation. Beyond sophisticated instruments and spacecraft, the intense radiation could also pose a threat to the safety of astronauts. Some basic devices that rely on the magnetic field, such as compass, would largely be distracted by geomagnetic explosions. The geomagnetic explosion caused by solar activities also leads to several natural phenomena and disasters, including drought, flood, earthquake, and thunderstorms. The high energy particle flows also create aurora by disturbing Earth's magnetic field.

As in the 21st century, communication systems and power supply are indispensable for humans. The emission of unanticipated solar flares would be catastrophic for those services. Therefore, building an accurate prediction model would be necessary for human decision-makers to avoid damage.

The aim of this paper is to develop a solar flare forecast model using MTF images, based on different machine learning models, including LeNet, AlexNet, VGG16, VGG19, ResNet, and Vision Transformer (ViT). Then, the effectiveness of each model will be tested. The major components in the designed experiment have been expressed in Figure 1. This paper contains the following sections: in section 2, the raw data undergoes preprocessing and forms MTF images; in section 3, the accuracies of the models are analyzed and results are presented; in section 4, the conclusion is presented with potential future

ISSN 2616-5775 Vol. 6, Issue 8: 158-164, DOI: 10.25236/AJCIS.2023.060819

improvements.

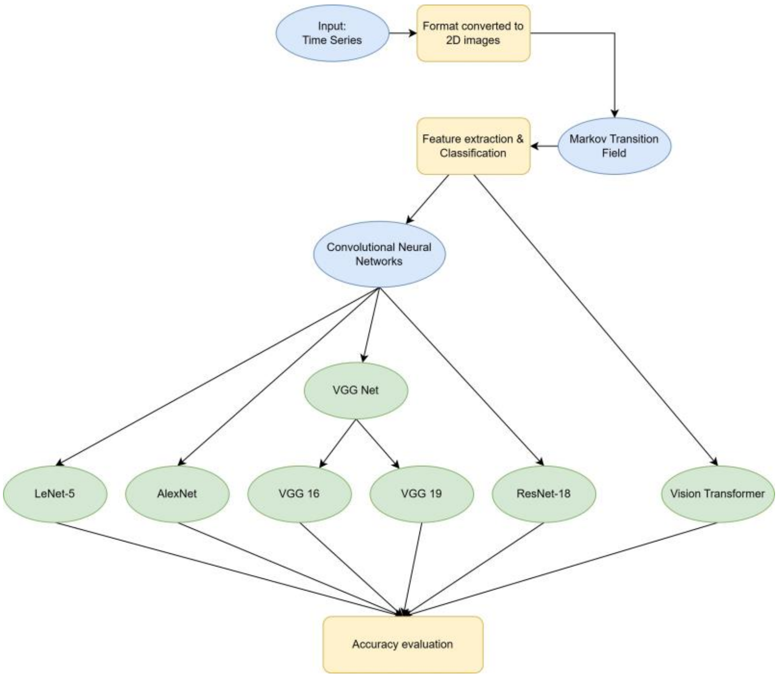

Figure 1: Project framework

2. Data preprocess

2.1. GOES-R XRS data

The data used in this paper is based on a database provided by NOAA (National Oceanic and Atmospheric Administration). The GOES-R satellites are responsible for monitoring weather conditions, including space weather that may cause disturbances to electrical systems on the surface of Earth. This database, GOES-R Level 2 Data, is de- veloped as the product of the GOES-R satellites' space weather monitoring. The selected data is provided by the 1-minute averages of XRS (X-Ray Sensor) of the measurements of EXIS (Extreme Ultraviolet and X-ray Sensors) of the GOES-16 satellite.

2.2. XRS time series data

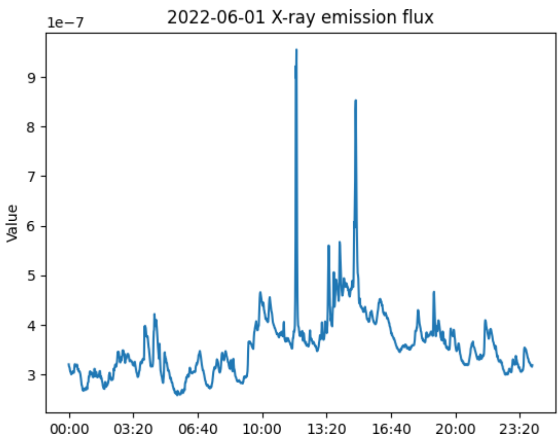

Figure 2: The initial time series data for Jun 3rd, 2022

All the data employed in this research are downloaded from the database of GOES-R. From the XRS 1-minute Averages files, a dataset of the continuous measurement of flux intensity is obtained. This dataset does not distinguish the occurrence of solar flare emis- sions, while the XRS Flare Summary files provide the moment of emission, flare classes, and integrated fluxes. However, while downloading files from the database, it is found that there exists massive deletion of data from specific months. Those missing months are not included in the collected dataset. At the time when the data were downloaded, there were only 2 months of data from 2020 recorded in the database. Therefore, the data utilized in this research includes years from 2017 to 2022, excluding 2020. In Figure 2, the initial time series data for Jun 3rd, 2022 has been illustrated. The horizontal axis shows time information and y vertical axis shows X-Ray flux[6-9].

2.3. Training dataset

As mentioned in subsection 2.2, the XRS Flare Summary enables the classification of data, distinguishing data with the occurrence of flares from those without. Two groups of MTF images used as the training data have developed. The first group has divided the dataset into two sets: Normal, where no flare is recorded in the 1 hour shown by the MTF; Flare, where flare data is matched to the Summary data in the MTF. The second group is divided differently: the MTF images that represent the solar activity 1 hour before the occurrence of solar flare have been extracted from the aggregated dataset; the rest of the data encompass both MTF that shows flare emission and not, while no MTF that records the hour before the flare is included.

According to the generalization ability of deep learning models, prediction can be performed even if the training data is not expressing the prior circumstance. However, training the model with flux data representing the solar flare simultaneously may render the reliability of the accuracy to be reduced. The feature extracted by the models learning the first data group can capture both the features prior to flare emission and the features while the flare is emitted; models learning the second data group would mostly capture the features prior to flare emission. Thus, both methods of model training could produce applicable models, however it is speculated that the second data group could produce more reliable models, regardless to the eventual accuracy.

Since a size of 60×60 is relatively small, the input image to the models has been resized to 224×224. The initial dataset contains a total of 81,504 MTF images, of which 60% of the dataset has been encompassed by the training set and the other 40% is in the testing set.

3. Results and analysis

After all models are fully trained, the loss functions have been plotted and the models are been applied to the testing set for obtaining the accuracy value.

3.1. Training results

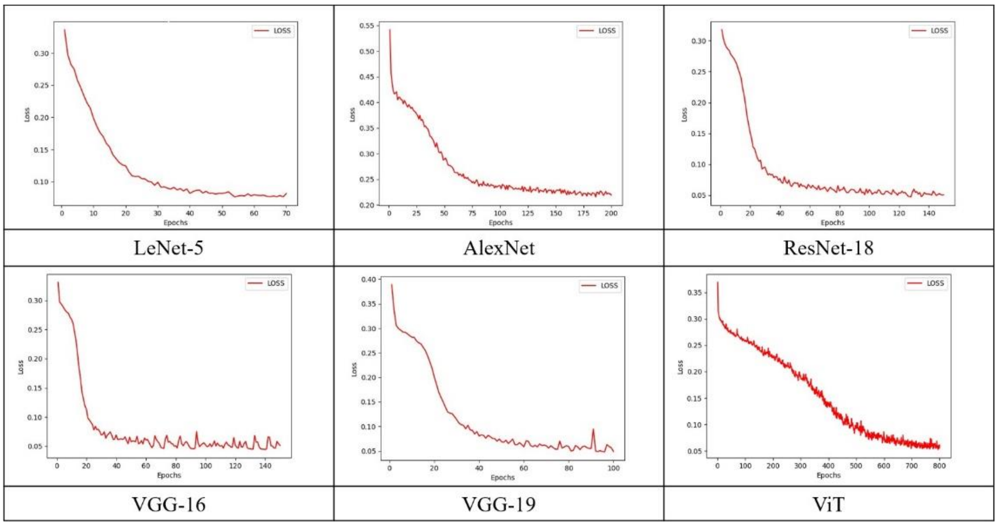

Figure 3: Loss function of the trained models with dataset 2

According to the different architects of models, different epoch numbers are required to fully train models. After each model is trained, the corresponding loss function is illustrated in Figure 3. The horizontal axis indicates the number of epochs trained and the vertical axis shows the loss. When the loss function approaches to stable and the tendency of descending is eliminated, the model is recognized as

ISSN 2616-5775 Vol. 6, Issue 8: 158-164, DOI: 10.25236/AJCIS.2023.060819

fully trained. As shown in the graph, LeNet-5 get enough training around 50 epochs; AlexNet shows enough training around 150 epochs; VGG-16 gain sufficient training around 40 epochs; VGG-19 get enough training at 60 epochs; ResNet-18 gain enough training around 60, and express minor decline afterward; loss function of ViT kept descending until around 700 epochs. However, since the data set of flare MTF images is self-made, some loss functions did not eliminate fluctuations even if reaching a stable tendency.

The results of models based on data group 1 is shown in Figure 4; however, since this group of models is unreliable and has the possibility of misinterpreting the information in MTF, its accuracy is not worthy of consideration. Thus, only the accuracies of the second group of models are shown in the next subsection.

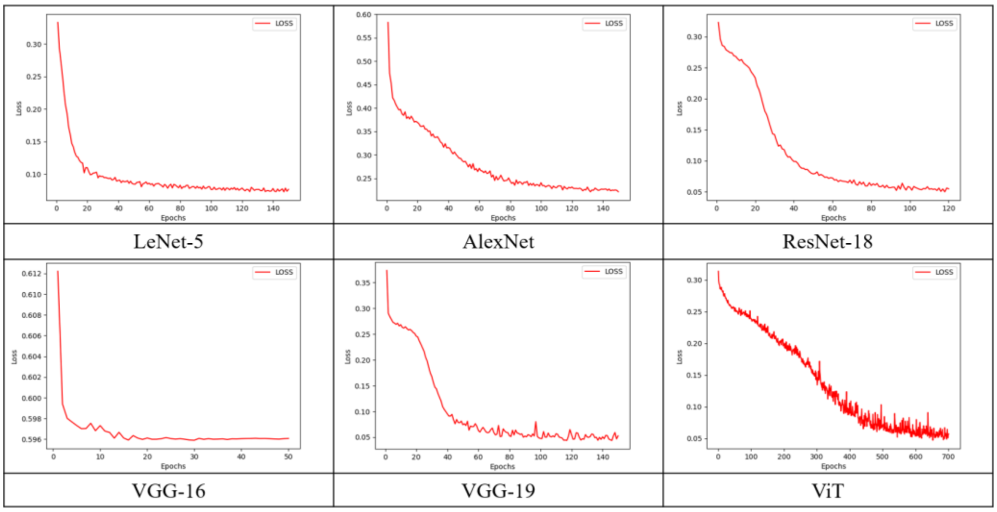

Figure 4: Loss function of the trained models with dataset 1

3.2. Accuracy

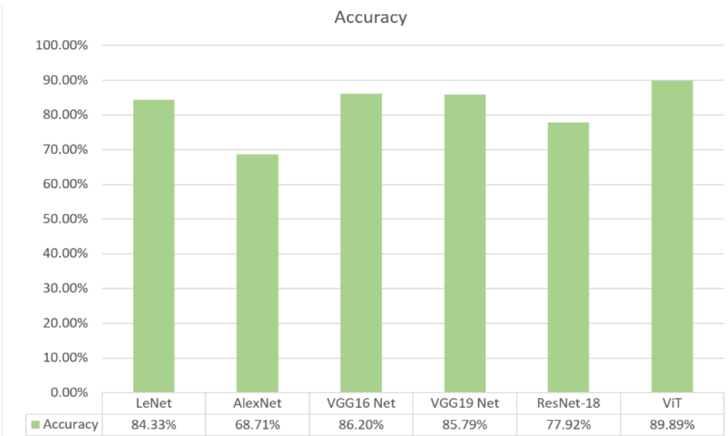

Figure 5: Accuracy of each trained deep learning model

The training set takes 60% of the dataset, hence the testing set contains the rest 40% of the data. Each model has been applied to the testing set and processed a classification work, eventually using the original label of the testing set to calculate the accuracy of the model. Accuracy is the percentage that expresses the counts by the classification compared to the labels in the testing set. As shown in the histogram and table in Figure 5, LeNet-5, VGG 16, and VGG 19 all have an accuracy greater than 80%; while Vision Transformer has the greatest accuracy level of 89.89%[10-11].

4. Conclusion

On the classification task of identifying MTF images that represents features of po- tential emission of solar flares, Vision Transformer shows outstanding performances and accuracy compared to other

ISSN 2616-5775 Vol. 6, Issue 8: 158-164, DOI: 10.25236/AJCIS.2023.060819

convolutional neural networks. Using the MTF translation method suggested in this paper, time series data of XRS X-Ray flux could also be utilized in solar flare forecasts with the use of deep learning models. Considering the outstand- ing performance of ViT (high accuracy of 89.89%) compared to other CNN models, ViT should be prioritized in forecast tasks of solar flare.

4.1. Advantages of methodology

Transforming raw data to imagery data is not innovative in the field of computer vision, however, converting our time series data to MTF form can retain the time in- formation in the images, which guaranteed the model could learn the features regarding time. Moreover, transformer is an architecture that was originally applied in the field of NLP, its application in computer vision, ViT, is a recent innovation that brought the accuracy of deep learning to a new height.

4.2. Potential improvements

There still exists a lot of work to be done in order to perfect this forecast model. Firstly, the major task of this paper is a binary classification problem, in which only ”Flare” and ”No flare” have been distinguished. To distinguish the solar flares with re- spect to A, B, C, M, and X levels, a multi-classification task has to be developed. This improved forecast model will have more applicability to real-life situations: since the decision-makers can distinguish the level of hazard of the flare to human civilization, degradation to instruments or communication systems can be minimized. Secondly, fur- ther correlation of the emission of flares of different levels with parameters other than X-Ray flux should be investigated. This could examine or even reinforce the reliability of utilizing X-Ray flux as the major parameter of this forecast task. Thirdly, only several prevalent CNN models are applied in the part involving CNN. Indeed, those networks have shown great accuracy and effectiveness when performing classification. However, the majority of applied networks do not encompass a variety of deepness. The effec- tiveness of some CNN may be improved through increasing layer numbers, hence further experiments could be designed. The structure of CNN could also be modified to adjust to optimize the performance in this experiment. Fourthly, similar to CNN architects, ViT could also undergo testing with varying structures, such as changing the head number or the convolution kernel. Resemble to the successful application of ViT contrasted by CNN, further latest models used in the computer vision field could be applied to this ex- periment. Fifthly, the extraction of the dataset could follow a more precise pattern. The current models in this paper are trained from the ”flare” labels indicating a solar flare or flares will emit in certain moments of the following hour. Therefore, the extracted fea- tures illustrate the general features of the time series t minutes before an emission, while $t \leqslant 60$. To improve the preciseness, the time series data of the exact 60 minutes before an emission could be compiled into an MTF image. By repeating this process instead of the prior segregation strategy in this paper, the randomness of the features could be further eliminated and the interpretability of the MTF images would enhance. Finally, applying the conversion of time series into 3-dimensional imagery information is not prevalent in the related field before. The effectiveness of this processing strategy in this certain field should be compared with other deep learning models, such as the CNN-LSTM model that distinguish active regions on an X-Ray photograph. The effectiveness of this paper’s methods can be examined and contrasted with deep learning networks based on other data forms. Furthermore, there exists a wide range of similar converting methods other than Markov Transition. Such as Fourier transformation, Recurrence Plots, and Gramian Angular Summation/Difference fields [12]. The applicability of those transition strategies could be reconsidered.

References

[1] A. Stanislavsky, M. M. K. B. J. J., W. Nitka. (2020). Prediction performance of hidden markov modelling for solar flares.

[2] Bo Wang, J. D., & Jiang, H. (2022). Markov transition field combined with convolu- tional neural network improved the predictive performance of near-infrared spectroscopy models for determination of aflatoxin b1 in maize.

[3] J. P. Mason, J. T. H. (2010). Testing automated solar flare forecasting with 13 years of michelson doppler imager magnetograms.

[4] N Nishizuka, Y. K. M. D. M. I., K Sugiura. (2018). Deep flare net (defn) model for solar flare prediction.

[5] Patricia L. Bornmann, J. H. L. M. R. G. H. A. G. R. V., David Speich. (1996). Goes x-ray sensor and

ISSN 2616-5775 Vol. 6, Issue 8: 158-164, DOI: 10.25236/AJCIS.2023.060819

its use in predicting solar-terrestrial disturbances.
[6] Wheatland, M. S. (2000). The origin of the solar flare waiting-time distribution. Wheatland, M. S. (2004). A bayesian approach to solar flare prediction.
[7] X Huang, L. X. J. L. R. L. X. D., H Wang. (2018). Deep learning based solar flare forecasting model. i. results for line-of-sight magnetograms.
[8] X Li, X. W. L. W., Y Zheng. (2020). Predicting solar flares using a novel deep convo- lutional neural network.
[9] Yanfang Zheng, X. W., Xuebao Li, & Zhou, T. (2019). Modified convolutional neural network with transfer learning for solar flare prediction.
[10] Yang Chen, A. O. H. G. T. B. D. T. Z. X. W. H. Z. Z. S. T. I. G., Ward B. Manchester. (2000). Identifying solar flare precursors using time series of sdo/hmi images and sharp parameters.
[11] ZhiguangWang, T. O. (2015). Imaging time-series to improve classification and impu- tation.
[12] Z Sun, X. W. Y. W. H. S. T. G. Y. C. A. H., MG Bobra. (2022). Predicting solar flares using cnn and lstm on two solar cycles of active region data.