Ubotica Technologies Ltd.
Innovation House
DCU Alpha
Old Finglas Road
Glasnevin
Dublin
D11 KXN4
Ireland

Title: **3CSD for Space Weather**

Document: **Executive Summary Report**

David Moloney PhD
Chief Scientist, Ubotica
1/12/2023

1 Introduction

This document represents an executive summary of the 3CSD proof of concept.

The 3CSD space weather features are mainly divided into four main features:

- Continuously monitoring the sun in real time to swiftly detect and track solar events like the evolution of the active regions and Coronal Mass Ejections (CMEs)
- EoFNets (Eye on Flare Networks) represents the stack of two-time series networks to predict whether the active region is flaring. The first network is EoFPhyNet which is a physically-based feature-detector that is developed to predict whether the active region is flaring using the physical features of the active regions. The second network is the EoFGeoNet which is designed to predict whether the active region is flaring using geometric features.
- FrCMENet (Flare-related CME Network) is a time series network that is designed to predict weather flares associated with CME using flare features. In addition, CMESP (CME Startdate Prediction) is designed as a time series network as part of FrCMENets to estimate the start date of CME using flare features.
- CMEPP (CME Properties Prediction) is dedicated to predicting the CME properties such as Central PA (deg), Angular Width (deg), and Linear speed (km/s) using LASCO C2 and LASCO C3 datasets.
- CMEATNets to predict the arrival time of CME that hits the Earth

The following sections highlight the main idea and proposals developed for the space weather use case.

2 Detection/Tracking

2.1 Real-time monitoring

YOLOv5 [10] and Deepsort [8] are employed for real-time solar events detection and monitoring, providing a trade-off between high accuracy and less memory usage. The dataset used for solar event detection is collected from the Solar Dynamics Observatory (SDO) website [3]. The data used in the detection/tracking workflow is real-time streaming [1]. A late fusion technique is applied to leverage multiple data sources, enhancing detection and tracking outcomes. Active regions are saved in FITS format for user requests and as time series input for networks like EoFPhyNet and EoFGeoNet. Notably, the results illustrate the enhanced detection of coronal holes through an ensemble approach compared to both single models and intermediate fusion methods [5] [6].

2.2 Multistage LILLIAN

To enhance the original model performance and ensure high detection accuracy, the Multistage LILLIAN approach is proposed. Multistage LILLIAN originates from the LILLIAN API which has been extended to support multiple inputs and multiple models. The qualitative results of the LILLIAN API confirm that the LILLIAN API can provide newly detected objects missed by the original models. The experiments confirmed that the Late Fusion (Voting System) surpasses Intermediate Fusion and aligns closely with the individual model results.

2.3 Multistage NNS

The NNS [9] serves as a monitoring tool, enabling runtime checks to determine whether a given pattern was seen by the network during its training phase. This is achieved by using a specific intermediate layer for the pattern analysis. In the context of multistage NNS, two distinct approaches, namely Intermediate Fusion and Late Fusion (Ensembles), were implemented and evaluated at the decision stage.

3 EoFNets

EoFNets is a stack of two-time series networks used to predict whether the active region is flaring or not. The EoFNets fuse the results of the two separate networks: EoFPhyNet and EoFGeoNet to provide fused results. EoFPhyNet is a physically-based features (SHARP) network designed to predict which active region is flaring. EoFGeoNet is designed to extract time series geometric features from time series active region patches using the GeoNet Extractor [4] and use these extracted features to predict whether the active region is flaring by analysing its geometric evolution over time. Figure 1 Highlights the main parts of EoFNets.

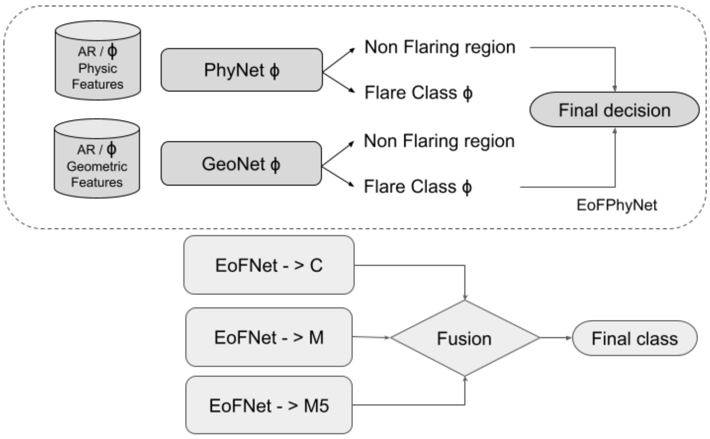

Figure 1: EoFNets workflow

4 FrCMENet

The FrCMENets is composed of the FrCMENet network which is a time series model to process the active region evolution and predict whether the processed active region is flaring and producing CME. CMESP is designed to process the solar flare time features in a specific window of time to estimate the CME start date that could be produced from the same Active region. The confusion matrix of the FrCMENet shows an error rate of 40% meaning around 40% of the total predictions made by the model are incorrect. However, only 15% of the Flare without CME class and Flare associated with CME class are predicted as a Non-Flaring region (class 0) by the model. The CMESP model occasionally exhibits a margin of error ranging from 2 to 3 days in some cases

5 CMEENets

We built the CMEATNets V1.0 which is a CME workflow to estimate and predict the CME features. The CMEENets is composed of two main parts: CMEPP which is designed to predict and estimate the CME features from the LASCO coronagraph image and the CMEAT which is the combination of two regression models. The first network is dedicated to predicting the arrival/shock time of CME using CME properties such as angular width, main position angle, linear speed, and mass.

6 Benchmarking

We conducted benchmarking of various solar event models across diverse low-power hardware platforms, including MyriadX VPU [7] and Rockchip 3588S[2]. Figure 2 represents the power consumption results of some 3CSD solar events models encompassing both detection and prediction tasks. Notably, the Rockchip power consumption is 2-3x higher than MyriadX (2-2.7x in performance efficiency OPs/W).

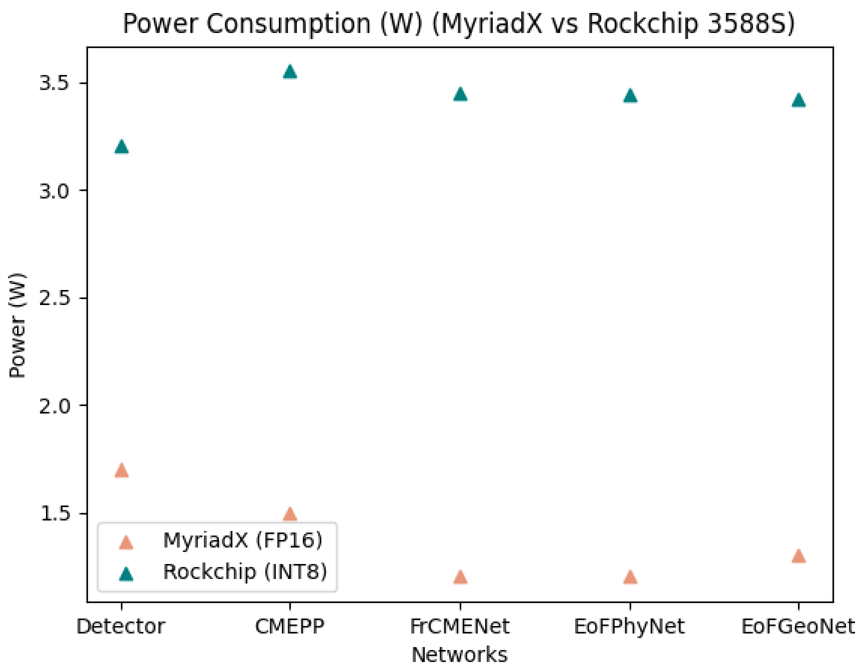

Figure 2: Power consumption (W)

7 Future work

In light of our current results and achievements, it is imperative to outline the plan for future work and the optimisation steps that will further enhance the developed systems. The following points describe our plans for continuous improvement and expansion beyond our current accomplishments in terms of space weather forecasting.

- The limitation of time series data will be handled by extending our GSSD (Get Synchronised Solar Data) tool to gather more data to enhance the capabilities of our time series predictors (EoFNets, FrCMENets, etc.)

- Accurate monitoring and prediction of CME features are still active topics for our work. One of the most important required features in our pipeline is the prediction of the arrival time of the CME that is directed toward the Earth.

- Through the CME properties prediction (CMEPP) that we have completed, we intend to extend the work by adding the prediction capabilities of other CME features to build a CME classifier based on the CME dynamic properties. For the Flare and CME early warnings (Prediction), the Active region tracking was used to control the evolution of the active region properties using HMI Magnetogram data. However, the sunspot properties could be also considered (e.g. using McIntosh classification) to track the sunspot properties' evolution over time and predict the active region class (Non-flaring region, Flaring region, Flaring region associated with CME).

- In the current study, the physics-based features and the geometric-based features of the active regions were considered. However, tracking the topology of Active regions also could be used to quantify the different changes happening[4]. The complexity of a magnetic field could be a clue to predict whether an active region will produce a flare, CME, or not. Thus computing the topology features of the Active region would be the next step to improve the performance of the prediction process.

- The possibility of retraining and updating the neural networks on board with fewer resources (power, memory, etc) using newly collected data, offers the advantage of real-time adaptability and continuous improvement. While retraining, a decrease of more than 30% in memory and time resources compared to those needed for the training process. This feature allows the solar events models to be updated to the changing conditions and enhances their decision-making capabilities by keeping them up to date and continually refining their performance over time.

- Some of the state-of-the-art retraining approaches are experiments, including incremental learning, which has the advantage of the training distribution over time which removes the CPU bottleneck during the training and continuously improves the performance of the model.

References

[1] https://cdaw.gsfc.nasa.gov.

[2] https://gadgetversus.com/processor/rockchip-rk3588-vs-rockchip-rk3588s/.

[3] https://sdo.gsfc.nasa.gov/data/aiahmi/.

[4] Varad Deshmukh, Thomas E. Berger, Elizabeth Bradley, and James D. Meiss. Leveraging the mathematics of shape for solar magnetic eruption prediction. Journal of Space Weather and Space Climate, 10:13, 2020.

[5] H. Li and X. J. Wu. Densefuse: A fusion approach to infrared and visible images, 2018.

[6] H. Li, X. J. Wu, and J. Kittler. Infrared and visible image fusion using a deep learning framework, 2018.

[7] Evangelos Petrongonas et al. Paralos: A scheduling & memory management framework for heterogeneous vpus. In 2021 24th Euromicro Conference on Digital System Design (DSD), pages 221–228. IEEE, 2021.

[8] J. Redmon and A. Farhadi. Yolo9000: Better, faster, stronger, 2017.

[9] A. Sahu, N. Vállez, R. Rodríguez-Bobada, M. Alhaddad, O. Moured, and G. Neugschwandtner. Application of the neural network dependability kit in real-world environments. arXiv preprint arXiv:2012.09602, 2020.

[10] A. Wong. Yolov5, 2020.

