3rd URSI AT-AP-RASC, Gran Canaria, 29 May – 3 June 2022

Automatic detection of e-Callisto solar radio bursts by Deep Neural Networks

Fernández Ruiz, Mario*(1); Bussons Gordo, Javier(2); Prieto Mateo, Manuel(1); and Monstein, Christian(3)
(1) Space Research Group. Universidad de Alcalá. Ctra. Madrid-Barcelona, km. 33.600. 28805 Alcalá de Henares, Spain
(2) Departamento de Física. Universidad de Murcia. 30071 Murcia, Spain
(3) Istituto Ricerche Solari (IRSOL), Università della Svizzera Italiana (USI), CH-6605 Locarno-Monti, Switzerland

Abstract

The aim of this work is to build a complete system based on deep neural networks for automated burst recognition in radio spectrograms delivered by ground-based solar observatories.

In this summary paper, the automatic system is described stage by stage and preliminary results for a sample observatory are presented.

1 Introduction

Continuous monitoring of solar activity at all wavelengths is a key tool for space weather prediction. Solar radio bursts and flaring episodes have an impact on Earth as they cause disturbances in electric power lines, satellite communication or GPS location.

For this purpose, solar event reports are compiled daily in near-real time including observations at radio frequencies (Figure 1), as well as in the X-ray and optical ranges, contributed by both ground-based and satellite-borne instruments from teams all around the world.

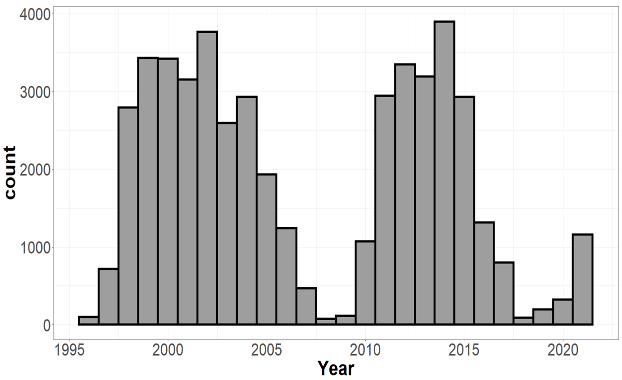

Figure 1. Number of solar radio events per year compiled by NOAA's Space Weather Prediction Center since 1996, showing solar activity cycles #23, 24 and the onset of 25.

The e-Callisto International Network of Solar Radio Spectrometers [1] comprises over a hundred instruments around the globe which altogether upload thousands of spectrograms per day to its central database. For that reason, manual classification of solar bursts is a challenging task. However, the number of articles in the field of automation of solar burst detection is still scarce.

Thus, in this summary paper we describe the stages of a complete automated system: pre-processing of spectrograms, conversion to images (PNG files), training of classification models based on convolutional neural networks and performance evaluation.

2 Pre-processing

2.1 Raw data

As mentioned before, our raw data are spectrograms and our goal is to find those in which a burst occurs. These spectrograms are FITS files (Flexible Image Transport System) which can be found on the e-Callisto website (http://www.e-callisto.org/Data/data.html). This database contains files since 2012.

Typically, these FITS files, or runs, have a resolution of 200 pixels in the vertical axis (corresponding to the range of frequencies in MHz, which is not the same for all observatories) and 3600 pixels in the horizontal axis (corresponding to time in seconds). Each run has a duration of 15 minutes, so every pixel spans 0.25 seconds. Finally, the intensity of the signal is represented by a color code.

2.2 Data processing

Our objective with pre-processing is to make the most meaningful features of the input spectrograms stand out for later training of the artificial neural network. Also, since these networks are computing systems inspired by biological networks in the human brain, we decided to turn the spectrograms into images in which we, as humans, are able to distinguish 'Bursts' from 'NotBursts'. Pre-processing is performed via Python scripts.

Firstly, in order to eliminate recurring interference and highlight transient phenomena, we apply background subtraction of the average intensity in every frequency channel. Next, we split the spectrogram into 15 different parts (one for each minute, Figure 2). This approach makes bursts and other phenomena easier to spot. Finally, we use the Python library `pyplot` to save the spectrograms as PNG files with a resolution of 256 $\times$ 256 pixels –the file type and resolution required by AlexNet [2] , the neural network used for training.

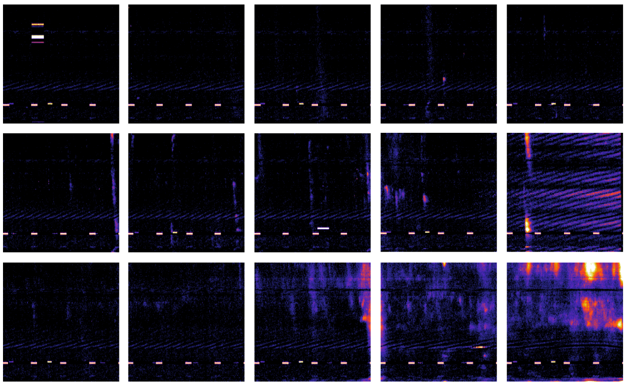

Figure 2. Cropping of an image in 15 parts – Glasgow, 8th November 2021, 13:00 to 13:15 UTC.

3 Training of classification models

We train neural networks using DIGITS: NVIDIA Deep Learning GPU Training System [3]. This is an interactive web environment with many different options to focus rapidly on training highly accurate deep neural networks (DNNs) for image classification [4].

Regarding the classification models, since the instruments in the e-Callisto network operate at different frequency ranges and each location has its own peculiar interference patterns, we currently build observatory-specific models where all the images entering the training dataset belong to the given observatory. So far we have built models for those instruments with a greater amount of recorded bursts, an information which is also available on the e-Callisto website. In the near future we intend to build models for groups of observatories with similar features.

All throughout the training and evaluation process, the truth of a prediction (True/False Positive/Negative) is assessed using the Callisto Event Reports available (http://soleil.i4ds.ch/solarradio/data/BurstLists/2010-yyyy_Monstein/). These are monthly 'Burst' lists compiled daily by an experienced analyst through visual inspection of each individual spectrogram image, a check of whether the burst has been detected by other e-Callisto stations and a cross-match with the corresponding NOAA Solar Event Report [5]. Years of careful, and tiring, inspection of thousands of these images have led to a good knowledge of the topology and features (in the time-frequency-intensity 3D space) of different types of burst signals and of the many types of fake signals caused by natural and man-made radio noise (see Catalog of Dynamic Electromagnetic Spectra [6]).

We train our models in two rounds. In the first one, an initial training dataset yields a base model which is then tested on a separate dataset for performance evaluation expressed via a so-called confusion matrix (see Tables 1 and 2).

This way we also identify the most common kinds of prediction errors, both as False Negatives (FN) and False Positives (FP). In the second round we use that information to try and improve performance by adding relevant images to the training dataset. We also mark the base model as pretrained to be used as a network for this second round, as pretrained models have better-distributed and quasi-optimized weights in their convolutional neural network layers, which saves computing time and allows for a better performance. This procedure allows us to improve the model significantly.

3.1 Training datasets

To train a model we need a dataset with a large amount of images of the two classes: Burst and NotBurst. The same Python script we use to pre-process the spectrograms is used to download all the FITS files of a given day and e-Callisto instrument and to convert them into PNG images.

In particular, for this summary paper we will discuss the results of a test on a model for the observatory in Glasgow. The initial training dataset contains images from May 22 and 24, 2021, two dates when the Sun was very active.

3.2 Training specifications

As it is well known, when training a classification model you need to split your data into a training and a testing set. Also, within the training set one part is used for validation. In our case we will use 5% for testing and 25% of the training set for validation.

Regarding the solver options in training, we set 100 training epochs and a validation interval of 10 epochs. Blob format is NVCaffe [7] and the solver type is SGD (Stochastic Gradient System) with a Base Learning Rate of 0.01. The neural network used is AlexNet.

4 Performance evaluation

4.1 Metrics for performance

In this section we will present the main metrics used to quantify the performance of the models. First, we introduce some concepts:

- **Positive (P)**: we denote the class Burst as positive.
- **Negative (N)**: the class NotBurst.
- **True positive (TP)**: predicted class is Burst; actual class is Burst.

- **False positive (FP)**: predicted class is Burst; actual class is NotBurst.
- **True negative (TN)**: predicted class is NotBurst; actual class is NotBurst.
- **False negative (FN)**: predicted class is NotBurst; actual class is Burst.
- **Confusion matrix**: a table for comparison of predictions vs actual classes.

From these concepts we compute some relevant metrics:

- **True positive rate (probability of detection):** ratio of true positives to the actual number of positives, $\frac{TP}{P}$.
- **True negative rate (specificity):** ratio of true negatives to actual negatives, $\frac{TN}{N}$.
- **False positive rate (probability of false detection):** ratio of false positives to actual negatives, $\frac{FP}{N}$.
- **False negative rate (probability of a miss):** ratio of false negatives to the actual number of positives, $\frac{FN}{P}$.
- **Accuracy:** $\frac{TP+TN}{P+N}$. This metric provides us with an overall idea of the performance.

4.2 Results

Results of the first-round test are presented in the following confusion matrix with the total numbers, and percentage rates in parentheses, of:
Top row: True Positives and False Negatives (red);
Bottom row: False Positives (yellow) and True Negatives.

Confusion matrix	Predicted Positives	Predicted Negatives
Actual Positives	21 (84%)	4 (16%)
Actual Negatives	56 (11%)	436 (89%)

Table 1. Total number (percentage rates) of TP, FN (red), FP (yellow) and TN.

As we can see, in the first round 84% of the actual positives listed in the reports were recognized, with a similar success rate (89%) for the actual negatives. The accuracy obtained is 88%.

There are four false negatives (16%) in total. A closer look at the missed bursts reveals, in one case, the end of a 20-minute-long storm which started and was correctly identified by the system in the previous run; and two cases of very weak or ill-defined signals classified as "hardly visible bursts" in the e-Callisto Event Report. If some or all of these doubtful false negatives were not counted, the FN rate would drop below 10%, even to 5%.

When it comes down to counting burst events, telling a bunch of individual bursts from a single storm is a delicate issue. Also, the start and end times of a given storm may be clear in those observatories where the signal is intense but hard to define in others; or they may be comnpletely different in different locations.

Regarding False Positives, their rate (11%) is still significant but from the point of view of a Quicklook Tool, the use of our system implies a huge reduction (of 89%) in the number of files to be inspected manually.

It is important to recall that this model is still in its first round. We are currently adding images belonging to common types of False Positives and False Negatives to the training dataset.

5 Conclusions

The burst classification model evaluated in this summary paper, despite not being definitive, gives significantly good results in terms of data reduction for visual inspection and promising results in terms of false negatives.

We are optimistic that this model will improve its performance in the second round, with a further reduction of False Positives and, more imortant, of False Negatives –work in progress, the results of which we expect to present by the time this conference takes place.

If this goal is achieved we intend to use this method to expand the e-Callisto Burst Database to the long period of time (years 2012–2019) for which FITS files are available but event lists are not. This would allow a thorough cross-match of e-Callisto events with the already-existing reports by NOAA.

Let us finish by mentioning a few ideas which can be explored in the near future. Once the presented model is brought through the second round, the same process will be repeated for other observatories. We should also investigate whether it is possible to build a model capable of recognizing bursts for several observatories with similar range of frequencies, interferences, etc. In addition, we are working on the full automation of the classification system so that anyone can use it. Finally, it would be very useful to train networks to classify bursts according to their type (I, II, III, CTM, ...).

6 Acknowledgements

This work is part of CELESTINA (Castillian E-callisto Leading Experimentation in Solar-Terrestrial Interaction with Novel Antennas), a project funded by the Junta de Comunidades de Castilla-La Mancha and the European Union (reference: SBPLY/19/180501/000237) with collaboration by Parque Científico-Tecnológico de Castilla - La Mancha.

We also thank Fachhochschule Nordwestschweiz FHNW, Institute for Data Science in Brugg/Windisch, Switzerland for hosting the e-Callisto central server and our ABSys-UCM/UEX collaborators for the use of DIGITS via the Micro-Stress- MAP-CM project (Y2018/NMT-4668).

References

[1] A. O. Benz and Ch. Monstein, “Callisto — A New Concept for Solar Radio Spectrometers,” *Solar Physics*, **226**, 1, January 2005, pp. 143–151, doi:10.1007/s11207-005-5688-9.

[2] Z. Alom, T. M. Taha, C. Yakopcic, S. Westberg, P. Sidike, S. Nasrin, B. C. Van Essen, A. A. S. Awwal, and V. K. Asari, “The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches”, arXiv:1803.0116v2 September 2018, pp. 1–11.

[3] NVIDIA, “DIGITS: Interactive Deep Learning GPU Training System”, updated 2022, https://developer.nvidia.com/digits .

[4] Y. Sun, X. Huang, D. Kroening, J. Sharp, M. Hill, R. Ashmore , “Testing Deep Neural Networks”, arXiv:1803.04792v4 September 2018.

[5] Space Weather Prediction Center (National Oceanic and Atmospheric Administration, USA), “Solar and Geophysical Event Reports”, updated 2022, https://www.swpc.noaa.gov/products/solar-and-geophysical-event-reports .

[6] Ch. Monstein, “Catalog of dynamic electromagnetic spectra”, updated 2020, http://www.e-callisto.org/GeneralDocuments/BurstCatalog.pdf.

[7] Berkeley AI Research (BAIR), “Caffe, a deep learning framework made with expression, speed, and modularity in mind”, updated 2022, https://caffe.berkeleyvision.org .