# DEEP LEARNING APPROACH TO CLASSIFYING SOLAR RADIO BURSTS

Herman le Roux, Günther Richard Drevin, Roelf Du Toit Strauss and Petrus Johannes Steyn
*North-West University Potchefstroom, South Africa*

**ABSTRACT**

Solar radio bursts have a profound effect on the environment and the technological infrastructure of society. Due to the effects that solar radio bursts have, it is important to be able to correctly classify them in an efficient and timely manner. Recently, there has been great success in object classifying by using convolutional neural networks. Therefore, this paper attempts to classify type II, and type III solar radio bursts, and spectrograms containing no bursts, using a convolutional neural network. An in-depth study was conducted to understand the nature of type II and type III bursts and how convolutional neural networks can be implemented to classify these bursts. Data samples, in the form of FITS files, were collected from the e-Callisto network and processed into images. A convolutional neural network was created with the Python packages TensorFlow and Keras. This model was then trained and tested using the only raw data collected from the e-Callisto network. While the model was not able to reach a very high accuracy the paper shows that such a model can make correct predictions.

**KEYWORDS**

Solar Radio Bursts, Space Weather, Deep learning, Convolutional Neural Networks

## 1. INTRODUCTION

Solar radio bursts have the potential to greatly affect our wireless networks on Earth as well as our near-space environment, where a lot of space exploration and missions occur. Solar radio bursts cause problems for cell phone towers and in some cases cause worldwide GPS navigation outages (Gary, 2011). Research shows that more than a billion devices are currently connected to the Internet of Things (Pahlavan and Krishnamurthy, 2020). This should cause great concern when one considers that a large enough burst can cause radio blackouts and severely damage wireless networks. Currently, the classification of solar radio bursts at e-Callisto is done manually. Christiaan Monstein, a retired Senior Radio Engineer working at IRSOL, creates event lists that contain information regarding which stations recorded different types of solar radio bursts and at what times, from the e-Callisto network. As one could imagine, it is quite a tedious task, as evidenced by the lack of radio burst lists on the e-Callisto website from 2012 to 2019 due to a lack of manpower. Solar radio bursts are very complex events and have very specific characteristics. In the following literature review, the necessary background information regarding the origin and causes of solar radio bursts will be discussed. To be able to create a neural network to make predictions, it is necessary to understand the different types of solar radio bursts as well as how they are observed using spectrometers. First, the magnetic Sun is described (1.1) followed by solar radio bursts (1.2).

### 1.1 The Magnetic Sun

In our solar system the Sun, a burning plasma ball of hydrogen and helium, is the main source of energy and space weather events. Most of the energy the Sun emits is electromagnetic radiation. The magnetic structure of the Sun's photosphere is constantly changing. The Sun's magnetic field varies in magnetic activity. Due to high magnetic activity, concentrations of strong magnetic fields will occur called sunspots (Hanslmeier, 2020). The solar cycle is an 11-year periodic variation in the number of sunspots that can be observed on the surface of the Sun. Due to the Sun's highly conductive plasma, a process called magnetic reconnection occurs regularly. Magnetic reconnection is the result of magnetic field lines in the plasma breaking and reattaching, causing magnetic energy to be converted into the heating and acceleration of particles (Hanslmeier, 2020).

Magnetic reconnection is a process that occurs when magnetic field lines stretch, wrap or twist building up tension and eventually breaking their current connection and reconnecting to a polar opposite magnetic field line close by (Hesse and Cassak, 2020). After these field lines reconnect, they straighten out which results in the release of energy in the surrounding region and in the case of the Sun; plasma is usually released as well (Hesse and Cassak, 2020). Magnetic reconnection also occurs in situations where magnetic fields change direction, which happens on the Sun quite often (Hesse and Cassak, 2020). This process is the reason for most explosive bursts of energy on the Sun, such as solar flares and coronal mass ejections (CMEs) (Hanslmeier, 2020).

## 1.2 Solar Radio Bursts

Events caused by magnetic reconnection such as solar flares and coronal mass ejections (CMEs) cause an increase in emissions in EUV (extreme ultraviolet)-, X-ray, gamma-ray, and radio waves from the Sun's already continuous spectrum (Bothmer *et al*., 2007). These solar radio bursts (SRBs) are observable using spectrometers and are usually either classified as Type II and Type III although there are five types of solar bursts which include: Type I, Type II, Type III, Type IV and Type V (Liu *et al*., 2018). We will mainly focus on Type II and Type III solar radio bursts in this paper.

### 1.2.1 Type II Bursts

Type II solar radio bursts occur after the maximum phase of flares which is commonly about 5-20 min after the start of the solar flare (Wild *et al*., 1963). It is accepted that these bursts originate in plasma waves (Wild *et al*., 1963). The outward movement at supersonic speeds suggests type II bursts are related to magnetohydrodynamic shocks (Wild *et al*., 1963). Strong shocks and shock-accelerated particles in the solar corona cause coronal mass ejections, which are events that are also highly associated with type II bursts (Cane and Stone, 1984). Coronal mass ejections are caused by magnetic reconnection which releases accelerated thermal electrons and plasma when they escape the reconnection site (Kou *et al*., 2020). The starting frequency of the meter wavelength type II burst, as seen in Figure 2, ranges from 70-100 MHz (Wild *et al*., 1963; Cane and Stone, 1984).

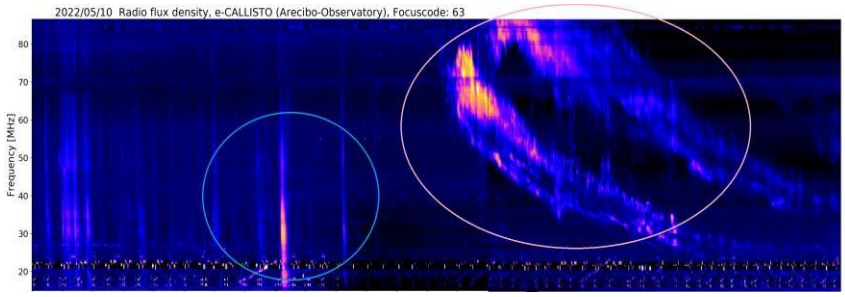


Figure 1. Type III and II solar radio bursts observed in Arecibo, Puerto Rico

### 1.2.2 Type III Bursts

Type III solar radio bursts individually last for approximately 10 seconds, yet they often appear to cluster in a group of ±10 as seen in Figure 1 (Wild *et al*., 1963). These bursts are strongly associated with the occurrenceof solar flares and usually occur at the start of these flares (Wild *et al*., 1963). These solar flares are the sudden intense events in the solar atmosphere that enhances electromagnetic radiation and accelerate electrons (Kou *et al*., 2020). Type III solar radio bursts drift at a rate of 2 to 16 MHz per second (Ndacyayisenga *et al*., 2021). The starting frequency for these meter wavelength bursts is 5 to 600 MHz (Wild *et al*., 1963).

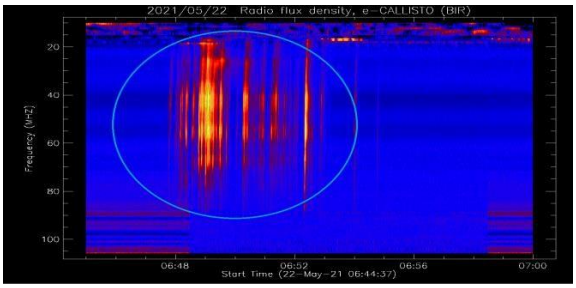


Figure 2. Type III solar radio bursts observed in Bir, Ireland

### 1.2.3 Summary of the Literature Review Regarding Solar Radio Bursts

The events produced through solar activity, including solar flares and coronal mass ejections, are products of the intense magnetic activity that is present on the surface of the Sun due to sunspots and the chaotic nature of the Sun's magnetic field lines. Solar radio bursts are spikes in radio waves caused by magnetic activity and within them lies the main cause of solar radio burst-related disturbances in our communication systems. The ability to forecast space weather events caused by these solar radio bursts is important due to these effects. When comparing Figure 1 and Figure 2 we can visually discern between the slower drifting type II and the faster drifting type III bursts.

## 2. CREATING THE DATASET

The e-CALLISTO network is a combination of all the data provided by CALLISTO instruments that are set up around the world. The spectrograms of solar bursts made by these instruments are plotted and stored as frequency over time graphs in FITS (Flexible Image Transport System) files, as seen in Figures 1 and 2. These stations are set up in a wide varietyof different ways but all of them provide the data with Coordinated Universal Time (UTC) timestamps. Data samples from January 2021 to August 2022 collected by the e-CALLISTO network were downloaded using thewebsite scraping library Beautifulsoup. These files were downloaded as compressed files and extracted to FITS files. The files were sorted into folders for type II and type III bursts as well as empty spectrograms using the labelled events list in the form of text files Christian Monstein provides. The FITS files were then divided randomly into training-, validation-, and test sets with the corresponding ratios of 0.7, 0.2 and 0.1. Due to type III bursts occurring more often than type II bursts stratified sampling was used to reduce empty- and type III samples to minimize the effect of data bias. While further refinement of data samples could have been done byensuring that the images recorded at different frequencies, due to different installations of the spectrometers, which can be observed at the end of file names e.g., HUMAIN_20220415_080000_**59,** ALASKA- HAARP_20220710_001500_**62** it was purposefully avoided in an attempt to ensure that the model would be able to predict raw data as is from any e-Callisto spectrometer. The final training set made use of 4690 samples,2088 of which being empty spectrograms. There are 1305 type II burst samples and 1297 type III bursts.

## 3. DEEP NEURAL NETWORKS

To create neural networks that make it possible to automate tasks, we need to understand how neural networks work. Deep neural networks (DNN) are a subset of machine learning and a tool that makes use of algorithms based on the human mind called artificial neural networks (Trask, 2019). One of the 5 types of neural networks is feedforward neural networks (FNN) (Shrestha and Mahmood, 2019). Convolutional neural networks (CNN) are implementations of FNNs and are considered responsible for the biggest leaps in our current understanding and implementations of feature detection and image classification (Shrestha and Mahmood, 2019; O'Mahony *et al*., 2019). In this section, we discuss convolutional neural networks (3.1) followed by the Performance of convolutional neural networks on Solar radio bursts (3.2).

## 3.1 Convolutional Neural Networks

CNN is a popular and preferred neural network architecture for computer vision and is based on the visual cortex systems of humans (Shrestha and Mahmood, 2019). CNNs make use of kernels, also called filters, to detect edges and objects in images by spatially convolving through an entire image (O'Mahony *et al*., 2019). The CNN then uses the dot product of the kernel and the area of the original image to calculate the probability of the existence of that feature (O'Mahony *et al*., 2019).

### 3.1.1 Convolutional Layers

Starting at the top left of an input matrix the convolutional kernels (frequently 2x2 or 3x3) traverse through input data and calculate the product of the numbers at the same location in the input and the overlapping kernel (Wu, 2017). By adding these products together, we obtain a single value (Wu, 2017). This process is repeated for each possible input value until the bottom right is reached. Different types of kernels and biased operators can be added to the equations to highlight specific objects in images (Wu, 2017).

### 3.1.2 Pooling Layers

Pooling layers are layers inserted between convolutional layers to achieve sub-sampling which reduces the amount of memory used by the CNN and also speeds up the training process (Shrestha and Mahmood, 2019). Average pooling and max pooling are the two most used techniques (Yu *et al.*, 2014). Average pooling is when the pooling operator maps a sub-region to its average and max pooling is when the pooling operator maps the subregion to the maximum value (Wu, 2017).

### 3.1.3 Fully Connected Layer

The fully connected layer then takes the input from feature extraction and pooling layers and uses weights to predict which labels to assign to the input images (Shrestha and Mahmood, 2019). By using activation functions, the output of the convolution layers can be summed with a bias term to facilitate the learning of these weights (O'Mahony *et al*., 2019). Rectified Linear unit (ReLu) is a commonly used activation function in CNNs due to its results in the field of image recognition (O'Mahony *et al*., 2019). ReLU's main goal is to make CNN more nonlinear (Wu, 2017). We want the mapping from CNNs input to its output to be very nonlinear because the semantic information in a picture is a highly nonlinear mapping of pixel values in the input (Wu, 2017). By using the functions' ability to activate features it is very useful when attempting to recognize complex patterns and objects in images (Wu, 2017). ReLU is also useful in stochastic gradient descent learning due to the function not suffering from the vanishing gradient problem (Wu, 2017). Therefore, the gradients of the activated features are easily back-propagated (Wu, 2017). Backpropagation is a popular learning rule and it is commonly used in supervised learning tasks (Werbos, 1990). Backpropagation is used in these supervised learning tasks for pattern recognition and fault diagnosis and simply put if we increase the weights and the error increases, we lower the weights and vice versa until we reach an error as close to zero as possible (Werbos, 1990).

## 3.2 Performance of Convolutional Neural Networks on Solar Radio Bursts

In the last few years, a few studies have implemented neural networks to classify or detect bursts. Some have used generated data and others have implemented LTSM models. No study could be found that specifically attempts to classify different types of bursts with raw data and rather focused on being able to detect bursts in images. Results from previous studies will be discussed in 3.2.1, and 3.2.2 followed by a summary in 3.2.3.

### 3.2.1 University of Alcalá

The first study that will be discussed was conducted by the space research group at the University of Alcalá (Mario F.R. *et al*, 2022). The goal of the paper was to create a neural network that automizes the recognition of solar radio bursts. This paper also makes use of data provided by the e-Callisto network and the burst lists provided by Christian Monstein. Their model was tested and trained with an initial dataset with images from the Glasgow observatory of events recorded from May 22 and May 24 in 2021 (Mario F.R. *et al*, 2022). The model had a true positive rate of 84% and provides some alternative to manually identifying the bursts (Mario F.R. *et al*, 2022). While this is a positive trajectory for detecting bursts in a dataset it does not address the problem of differentiating between different types of bursts.

### 3.2.2 Laboratory for Electromagnetic Detection (LEAD), Institute of Space Sciences, Shandong University, Weihai, Shandong, China

This second study, published in 2022, was done at Shandong University and focused on creating a model that could quickly and automatically classify data between bursts and non-bursts (Guo J.C. *et al.*, 2022). Their model achieved an average classification accuracy of 98.73% (Guo J.C. *et al.*, 2022). Their paper also compares the performance of their model to previous research conducted relating to the same topic. It is clear that their paper greatly improves from an early model with a 67.4% true positive rate, with regards to burst detection, to 99.9% (Guo J.C. *et al.*, 2022). This study does not address the problem of differentiating between different types of bursts as well.

### 3.2.3 Summary of Literature Regarding Previously Created Neural Networks

Type II and type III solar radio bursts are a danger to modern networking infrastructures and being able to automatically detect and classify them are important efforts as they will add to astronomical research. When one considers the papers published and publicly available data on the e-Callisto website is clear that there exists a large corpus of data which is mostly manually labelled. Convolutional neural networks are a very prominent implementation of machine learning in computer vision, object detection and image classification and have been used with success by various papers for bursts identification.

# 4. PROCESSING THE DATA AND THE STRUCTURE OF THE NEURAL NETWORK

After the FITS files were sorted into the training-, validation- and test sets, several pre-processing steps occurred. In the following section the various techniques that were applied to each of the FITS files as well as the layers of the convolutional neural network created using the available Python libraries will be described.

## 4.1 Converting FITS Files to Images

A python library called radiospectra was used to import the FITS files. The frequencies were linearized and background noise subtraction was applied using the matplotlib library. The resulting graph was then plotted without the X and Y axis labels as shown in Figure 3. The plot function applied a Vmin equal to 5 to normalize the colour map gnist_ncar which makes use of 128 distinguishable colours on the graph. The imageswere then processed into grey images using a function from the pillow library that only stores the luminance of a pixel, effectively greyscaling the image. This was done to simplify the input shape to 1 channel of grey scale values ranging from 0-255 and avoid having 3 channels (Red, Green, Blue) as part of our input shape The image was then cropped to contain only the data which resulted in an output image of the size 324x357 pixels. Only the data was cropped from the graph because it is unnecessary to train the model with input featureslike the title of the observation and the intensity-colour scale. These images were used to train the model.

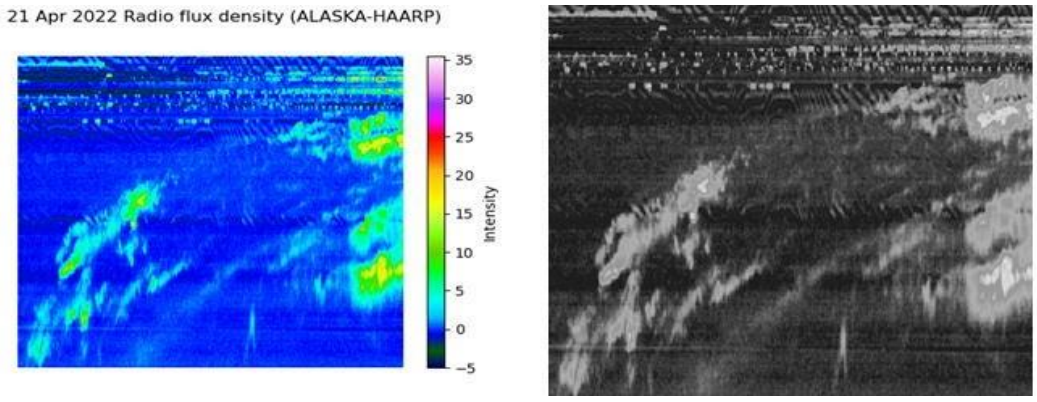


Figure 3. Displaying what the FITS files look like pre- and post-processing for a type II solar radio burst from the Haarp-Alaska observatory

## 4.2 The Structure of the Model

The TensorFlow flow_from_directory method was employed to load the data and fit the model. Due to the complexity of the images several convolutional layers were added to extract the necessary features to identify a type II and type III solar radio burst. Max pooling was applied to reduce the resulting output shapes. Four dropout layers were also added to prevent overfitting in the training set. The ReLu activation function was used in all the convolutional layers and the output was normalized using the SoftMax activation function.
Finally, the model was compiled using the sparse categorical cross entropy loss function and the adam optimizer function. The structure of the model is as follows:

Table 1. Structure of the convolutional neural network

| Type of Layer | Kernel size | Trainable parameters | Output shape |
|---|---|---|---|
| Input | N/A | N/A | (324,357,1) |
| Convolution2D | (3,3) | 160 | (322,355,16) |
| Maxpooling2D | (2,2) | 0 | (161,177,16) |
| Convolution2D | (2,2) | 1040 | (160,176,16) |
| Maxpooling2D | (2,2) | 0 | (80,88,16) |
| Convolution2D | (3,3) | 4640 | (78,86,32) |
| Maxpooling2D | (2,2) | 0 | (39,43,32) |
| Dropout (0.5) | N/A | 0 | (39,43,32) |
| Convolution2D | (3,3) | 9248 | (37,41,32) |
| Maxpooling2D | (2,2) | 0 | (18,20,32) |
| Dropout (0.5) | N/A | 0 | (18,20,32) |
| Convolution2D | (3,3) | 18496 | (16,18,64) |
| Maxpooling2D | (2,2) | 0 | (8,9,64) |
| Dropout (0.5) | N/A | 0 | (8,9,64) |
| Flatten | N/A | 0 | (4608) |
| Full connection | N/A | 4719616 | (1024) |
| Dropout (0.5) | N/A | 0 | (1024) |
| Output (Softmax) | N/A | 3075 | (3) |

# 5. EXPERIMENTAL RESULTS

The evaluation criterion (5.1.1) will be discussed followed by the results (5.1.2) and a discussion of the results (5.1.3).

## 5.1 Evaluation Criterion

Given that we need to create a confusion matrix for multiple classes we will need to create a multiclass confusion matrix to assess the experimental result. The true positive rate (TPR) is calculated as follows:

$$\mathrm{TPR} = \frac{TP}{\frac{TP+TF}{N}} \times 100\% \quad (1)$$

where TP denotes a correctly classed positive sample (true positive), and TFN denotes the total incorrectly classified positive samples. In general, a larger TPR value for a particular type of sample indicates a higher likelihood of this type's successful identification. The overall accuracy will be given as:

$$\mathrm{Accuracy} = \sum \mathrm{TP} / \sum \mathrm{Samples\ per\ class} \quad (2)$$

The precision for each class is determined as the TP of the class divided by the number of actual samples forthe class. The recall is calculated as the TP of the class divided by the number of samples predicted as that class.

## 5.2 Results

The model received a total number of 674 labelled samples for training with 299 being empty spectrograms, 188 being Type II solar radio bursts and 187 being Type III solar radio bursts the following confusion matrix was created:

Table 2. Multiple class confusion matrix detailing the results

| | Predicted Empty Spectrograms | Predicted Type II bursts | Predicted Type III bursts | Recall |
|---|---|---|---|---|
| Empty spectrograms | 270 | 12 | 17 | 90,30% |
| Type II bursts | 27 | 113 | 48 | 60,10% |
| Type III bursts | 24 | 34 | 129 | 68,98% |
| Precision | 84,11% | 71,06% | 66,49% | |

The diagonal indicates the true positive result for each class e.g., Type II bursts and Predicted Type II bursts intersect with a value of 113 indicating that the number of true positives for the class Type II bursts is 113. The overall accuracy of the model is 75,96%, the average recall is 73,13%, and the average precision is 73,89%. The recall is the number of samples correctly predicted as the class it was indicated as, and precisionis the number of samples correctly predicted from all the samples that were predicted as that class.

## 5.3 Discussion

The total number of samples containing bursts from the dataset was 20814 of which 18947 samples were type III solar radio bursts. The recall for type II solar radio bursts was 60,10% and the recall for type III solar radio bursts was 68,98%. Due to the number of type III bursts not being fully utilized to train the model in anattempt to minimize the data bias one could argue that the type III bursts recall could have been improved upon. The model struggled to correctly identify type III solar radio bursts and the model struggled even moreto identify type II solar radio bursts, but as can be seen in Figure 4, there exist samples in the testing set that have features that appear to be of type III solar radio bursts. The lack of type II burst data samples is prescribed to the fact that it is a rarer occurrence than that of a type III solar radio burst. This, accompanied bythe fact that some testing samples and training samples could be considered contaminated with type III solar radio bursts and other radio noise, is the reason for the lack of performance when considering the type II solarradio burst recall. The model performed well when predicting empty spectrograms with a recall rate of 90,30% but this can also be improved by using the entire corpus of data that accounts for empty spectrogramswhich in February of 2021 accounted for up to 89000 samples for the month. The average precision is 73,88% which indicates that, on average, the model predicted the true positives for a specifically correctly 73,88% of the time. This is encouraging when one considers that as 2022 progresses into a more active part of the solar cycle and more samples will be able to be collected and added to the training set. Further steps can be taken to improve the accuracy by filtering the data set such that type II and type III samples are extremely clear for human perception but this will rarely be the case and models should be focused on learning the features of the solar radio bursts despite massive amounts of noise of different causes.

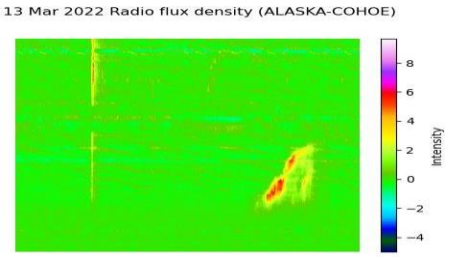


Figure 4. A type II solar radio burst from the test set was predicted as a type III solar radio burst

## 6. SUMMARY AND CONCLUSION

This paper is the first attempt at utilizing convolutional neural networks to automatically classify type II and type III solar radio bursts using a dataset constructed out of data downloaded between 2021 and 2022. A literature study was conducted to understand the nature of solar radio bursts, focusing on type II and type III solar radio bursts. Thereafter a literature study was conducted to understand convolutional neural networks and how the layers are utilized in terms of processing images and detecting features. A dataset was created using data collected from the e-Callisto website and event lists compiled by Christian Monstein. The data set was used to train a convolutional neural network created using the python libraries TensorFlow and Keras. The results of the model's prediction were given and discussed. The preliminary results have not shown a high accuracy but feature research is promising regarding the models' ability to identify the features of solar radio bursts.

## ACKNOWLEDGEMENT

The Institute for Data Science FHNW Brugg/Windisch, Switzerland is acknowledged for making the data publicly available as data from different years were used. Christian Monstein is also acknowledged for compiling the event lists that were used to create the dataset.

## REFERENCES

Bothmer, V., Zhukov, A., and Daglis, I. (2007). *The Sun as the prime source of space weather*, pages 31–102. Springer.

Cane, H. and Stone, R. (1984). Type II solar radio bursts, interplanetary shocks, and energetic particle events.

*The Astrophysical Journal*, 282:339–344.

Gary, D. E. (2011). Solar radio burst effects on wireless systems. In 2011 IEEE I*nternational Symposium on Electromagnetic Compatibility*, pages 661–664.

Guo J.C., Yan F-B, Wan G, Hu X-J, Wang S. 2022. A deep learning method for the recognition of solar radioburstspectrum. PeerJ Comput. Sci. 8:e855 DOI 10.7717/peerj-cs.855

Hanslmeier, A. (2020). *The Chaotic Solar Cycle*. Springer Singapore.

Hesse, M. and Cassak, P. A. (2020). Magnetic reconnection in the space sciences: Past, present, and future.

*Journal of Geophysical Research: Space Physics*, 125(2).

Kou, Y. K., Jing, Z. C., Cheng, X., Pan, W. Q., Liu, Y., Li, C., and Ding, M. D. (2020). What determines solar flares producing interplanetary type III radio bursts? *The Astrophysical Journal*, 898(1): L24.

Liu, H., Chen, Y., Cho, K., Feng, S., Vasanth, V., Koval, A., Du, G., Wu, Z., and Li,C. (2018). A solarstationary type IV radio burst and its radiation mechanism. *Solar Physics*, 293(4).

Mario F.R., Javier B.G., Manuel P.M. and Monstein C., Automatic detection of e-Callisto solar radio bursts by Deep Neural Networks, *2022 3rd URSI Atlantic and Asia Pacific Radio Science Meeting (AT-AP-RASC)*, 2022, pp. 1-4, DOI: 10.23919/AT-AP-RASC54737.2022.9814298.

Ndacyayisenga, T., Uwamahoro, J., Sasikumar Raja, K., and Monstein, C. (2021). A statistical study of solarradio type III bursts and space weather implication. *Advances in Space Research*, 67(4):1425–1435.

O'Mahony, N., Campbell, S., Carvalho, A., Harapanahalli, S., Hernandez, G. V., Krpalkova, L., Riordan, D.,and Walsh, J. (2019). Deep learning vs. traditional computer vision. In *Science and information conference*, pages 128–144. Springer.

Pahlavan, K. and Krishnamurthy, P. (2020). Evolution and impact of wi-fi technology and applications: Ahistorical perspective. *International Journal of Wireless Information Networks*, 28(1):3–19.

Shrestha, A. and Mahmood, A. (2019). Review of deep learning algorithms and architectures. *IEEE* Access,7:53040–53065.

Trask, A. W. (2019). *Grokking deep learning*. Manning.

Werbos, P. J. (1990). Backpropagation through time: what it does and how to do it. *Proceedings of the IEEE*,78(10): 1550–1560.

Wild, J. P., Smerd, S. F., and Weiss, A. A. (1963). Solar bursts. *Annual Review of Astronomy andAstrophysics*, 1(1): 291–366.

Wu, J. (2017). Introduction to convolutional neural networks. *National Key Lab for Novel SoftwareTechnology*. Nanjing University. China, 5(23):495.

Yu, D., Wang, H., Chen, P., and Wei, Z. (2014). Mixed pooling for convolutional neural networks. In *RoughSets and Knowledge Technology*, pages 364–375.