

# Applying saliency-map analysis in searches for pulsars and fast radio bursts

C. Zhang<sup>1,2,3,4</sup>, C. Wang<sup>3</sup>, G. Hobbs<sup>4</sup>, C. J. Russell<sup>5</sup>, D. Li<sup>6,2,7</sup>, S.-B. Zhang<sup>2,4,8</sup>, S. Dai<sup>4</sup>, J.-W. Wu<sup>1</sup>, Z.-C. Pan<sup>1</sup>, W.-W. Zhu<sup>1</sup>, L. Toomey<sup>4</sup>, and Z.-Y. Ren<sup>1</sup>

<sup>1</sup> National Astronomical Observatories, Chinese Academy of Sciences, A20 Datun Road, Chaoyang District, Beijing 100101, PR China

<sup>2</sup> University of Chinese Academy of Sciences, Beijing 100049, PR China  
e-mail: zhangchao215@mailsucas.ac.cn

<sup>3</sup> CSIRO Data61, Sydney, NSW 2015, Australia

<sup>4</sup> CSIRO Astronomy and Space Science, Australia Telescope National Facility, Box 76, Epping, NSW 1710, Australia

<sup>5</sup> CSIRO Scientific Computing, Sydney, NSW 2015, Australia

<sup>6</sup> CAS Key Laboratory of FAST, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, PR China

<sup>7</sup> NAOC-UKZN Computational Astrophysics Centre (NUCAC), University of KwaZulu-Natal, Durban 4000, South Africa

<sup>8</sup> Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210008, PR China

Received 2 December 2019 / Accepted 22 May 2020

## ABSTRACT

**Context.** We investigate the use of saliency-map analysis to aid in searches for transient signals, such as fast radio bursts and individual pulses from radio pulsars.

**Aims.** Our aim is to demonstrate that saliency maps provide the means to understand predictions from machine learning algorithms and can be implemented in pipelines used to search for transient events.

**Methods.** We implemented a new deep learning methodology to predict whether any segment of the data contains a transient event. The algorithm was trained using real and simulated data sets. We demonstrate that the algorithm is able to identify such events. The output results are visually analysed via the use of saliency maps.

**Results.** We find that saliency maps can produce an enhanced image of any transient feature without the need for de-dispersion or removal of radio frequency interference. The maps can be used to understand which features in the image were used in making the machine learning decision and to visualise the transient event. Even though the algorithm reported here was developed to demonstrate saliency-map analysis, we have detected a single burst event, in archival data, with dispersion measure of  $41 \text{ cm}^{-3} \text{ pc}$  that is not associated with any currently known pulsar.

**Key words.** methods: data analysis – techniques: image processing – methods: statistical – methods: numerical – pulsars: general

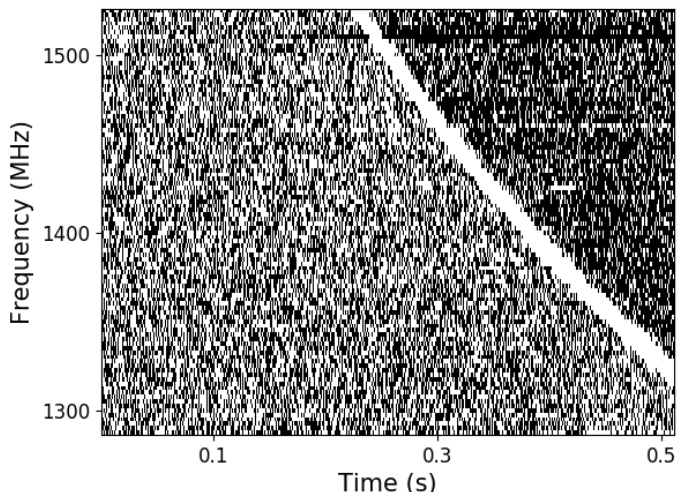
## 1. Introduction

Radio telescope observing systems continue to be used to record high time resolution data sets. In such data sets the total intensity of the received radio signal is sampled typically every  $\sim 100 \mu\text{s}$  and with moderate (e.g.  $\sim \text{MHz}$ ) channel bandwidths. For most historical data sets the samples are 1 or 2 bit digitised, but for many current surveys higher-bit data streams are recorded. High time resolution data sets are used to search for pulsars by seeking for weak periodic signals within the data. They are also used to search for bright, individual pulses from pulsars and fast radio bursts (FRBs; Devine et al. 2016; Michilli et al. 2018; Farah et al. 2019; Barsdell et al. 2012; Connor & van Leeuwen 2018).

Fast radio bursts are bright, millisecond-duration radio transients. The observed pulses are characterised by dispersion measures (DMs) that are significantly larger than the expected Milky Way contribution. They have been detected at flux densities between tens of micro-janskys and tens of janskys (Lorimer et al. 2007; Spitler et al. 2016; Connor & van Leeuwen 2018; CHIME/FRB Collaboration 2019a). Understanding the origin of FRBs is still an active research area, with many dif-

ferent theoretical explanations (Platts et al. 2019). The first FRB was discovered by Lorimer et al. (2007) during the reprocessing of archival pulsar survey data, and it is now commonly referred to as the “Lorimer burst”. A small segment of the data stream that was used to discover the “Lorimer burst” is shown in Fig. 1. This particular data file has 96 frequency channels (shown on the y-axis) spanning a total bandwidth of 288 MHz and each time-frequency sample (also referred to here as a pixel) has been 1 bit sampled.

Most of the known FRBs have only been detected once. However, there is now a small population of FRBs in which repeating signals have been detected (Spitler et al. 2016; CHIME/FRB Collaboration 2019a,b). There are likely thousands of detectable events each day across the full sky, but only a relatively small number have been published to date because of the moderate field of view that many radio telescopes have. Wide field of view telescopes such as The Canadian Hydrogen Intensity Mapping Experiment and the Australian Square Kilometre Array Pathfinder are now operating and a large number of FRB events will soon be published. However, for each detected FRB event, current survey processing methods usually produce thousands of false-positive triggers



**Fig. 1.** Part of the data file containing the original FRB event: the Lorimer burst. Each sample is 1 bit sampled, with white positive and black negative.

(Connor & van Leeuwen 2018). Some of these candidates can be rejected based on extra information, such as detection in multiple observing beams, but many of the diagnostic plots are simply inspected visually (Connor & van Leeuwen 2018). Future telescopes, such as the Square Kilometre Array, will carry out pulsar and FRB searches, but the enormous data rate from those telescopes implies that real-time processing is likely to be required. Real-time processing methods already operate for FRB searches (e.g. Barsdell et al. 2012), but produce large numbers of candidates (most of which are false-positive candidates).

Machine learning algorithms are increasingly taking a role in deciding which signals to record for further analysis. Ways to minimise their false positive rates and to maximise their efficiency and their robustness in the presence of radio frequency interference (RFI) need to be explored. Michilli et al. (2018) designed a machine-learning classifier to identify single pulses in a strong RFI environment, which relies on features such as the pulse width, DM, and signal-to-noise ratio (S/N) and has been used to discover seven pulsars. Farah et al. (2019) detected five new FRBs in real time with the Molonglo Radio Telescope. The pipeline adds an additional stage to the HEIM-DALL pipeline (Barsdell et al. 2012) to classify the resulting candidates using features extracted from time-frequency data. Connor & van Leeuwen (2018) focused on reducing the false positive rate from candidates obtained using a traditional search method. They applied a deep learning method to single pulse classification and developed a hierarchical framework for ranking events by their probability of being true astrophysical transients. Zhang et al. (2018a) presented the first successful application of deep learning to direct detection of fast radio transient signals in raw frequency-time data. They found 72 new pulses from the repeating fast radio burst FRB 121102 using the Green Bank Telescope.

All these algorithms report on whether a particular candidate (or image segment) contains a likely astrophysical burst event. However, they generally do not provide information on what parameters, or what features in the image were used to make that decision, and so deep learning models are often criticized as “black boxes” for decision making. This is primarily because the results from non-linear fitting of the high dimensional data is difficult to explain intuitively. A pulse verification process is often

needed to identify real signals from candidates produced by the machine learning modules. Many of these processes use visual inspection and assume that signals can be visually detectable directly by humans (Zhang et al. 2018a). In the presence of RFI or complex noise patterns, direct visual inspection is likely to miss real signals. We address this problem by ensuring that the machine learning procedure not only predicts whether an event is present, but also provides information on how it came to that decision. This line of work is categorised as machine learning interpretability.

There have been many efforts to address the problem of understanding why a machine learning method has made a decision, and to provide users with more confidence in the model predictions (see Montavon et al. 2018). For our purpose, we wish to identify the part of an input image that a classifier has identified as being from an astrophysical burst event and we make use of saliency analysis for this purpose.

Saliency-map analysis has been used for numerous ML applications to highlight features in an input that are relevant to the predictions of a model (Simonyan et al. 2013; Sundararajan et al. 2017), but to the best of our knowledge it has not been used on high-time resolution radio astronomy data sets to render the features of transient events.

The work described here is explicitly linked to the saliency-map analysis, and can be applied to the output of any deep learning method used to search for transient events. Of course, to demonstrate saliency-analysis we require a deep learning method and test data sets, but we note that saliency analysis is generally applicable and we are not attempting to convince the reader that our method is better or worse than any other existing algorithms for the FRB searching. In Sect. 2 we describe our machine learning algorithm, training procedure, and the test data set that we used for this work. We describe the saliency analysis method in Sect. 3, and conclude in Sect. 4.

## 2. FRB classifier and data set

### 2.1. Data set

To demonstrate the saliency analysis we make use of archival data sets that are publically available from the Parkes data archive<sup>1</sup> (Hobbs et al. 2011). The data sets are used to train the algorithm and to provide example observations (containing known pulsars, FRBs, and radio frequency interference) to test the effectiveness of the procedure.

All data sets were obtained with the Parkes telescope using the 21 cm multibeam receiver. The primary goal for carrying out the original observations was to search for new pulsars (Manchester et al. 2006). The data files are in PSR-FITS (Hotan et al. 2004) search mode format. They are two-dimensional spectrograms (time versus frequency) that span a frequency range from 1231 to 1516 MHz with 96 frequency channels. The time sampling varies between the different data files (from 125  $\mu$ s to 1 ms). The archive contains more than 100 observing projects, with each observing semester for each project stored as a data collection. Approximately 600 such data collections are now available for public access. Even though the observations were processed by the original science teams, new discoveries are still being made based on these archival data sets (recent discoveries have been reported by Pan et al. 2016; Zhang et al. 2018b, 2019).

A single observation often contains millions of time samples. We cannot simply pass the entire data file as an image into

<sup>1</sup> <https://data.csiro.au>

a machine learning classifier as typical algorithms require much smaller image sizes and the signals of interest (i.e. the astronomical bursts) only last for very short time durations and hence make up a tiny fraction of the entire observation. We therefore take each observation and split the file into small segments (for this work we use segments of 512 time samples). We divide the segments into two categories:

1. segments containing a burst candidate;
2. segments not containing a burst candidate.

A given data set may include receiver noise, RFI, bright individual pulses from pulsars, FRBs and other unexpected signatures. RFI usually takes the form of wide-band impulsive signals or narrow-band persistent signals, but can also mimic astronomical signals (Petroff et al. 2015; Men et al. 2019).

The classifier requires a training procedure. Unfortunately, the number of known FRBs in the Parkes data sets is relatively low and single pulses from known pulsars all have relatively small DM values. This implies that we cannot simply train the algorithm on actual signals in the archival data. Instead, we inject simulated burst events into 1000 randomly chosen data files from the Parkes data archive. We simulated the bursts assuming the frequency-squared dispersion law. The FRB event is therefore parametrised by a time (corresponding to the arrival time of burst at the highest observing frequency), the DM, a width (the FRB is assumed to have a Gaussian profile)<sup>2</sup>, and a brightness. As we are injecting simulated FRBs into 1 bit data we need to ensure that we have a way to simulate different FRB brightnesses. To do this we define the fraction of samples within the FRB envelope that will become 1 (representing a signal above the mean level) and how many will remain 0 (a signal below the mean level). We note that no value that is already 1 will become a 0 in this process to ensure that any existing signal, such as RFI, is not affected by the simulation process.

For our training data set we simulated a wide range of possible FRB parameters. The start time of the FRB was randomly chosen anywhere within the observation span, the DM values ranged between 20 and 5000 pc cm<sup>-3</sup>, the saturation level measured by the percentage of pixels within the signal range turned bright by the FRB ranged from 75% to 100%, and the width of the FRB was chosen between 3 and 50 time samples.

We randomly selected 1000 files from data collections 1 to 3 in Table 1, from which we extracted 57 000 data segments. We injected the simulated FRBs into 24 500 data segments. These resulting files therefore contained the real noise signals and our simulated FRBs. We have ensured that these specific data segments do not contain known FRBs or pulsars, although they may contain currently unknown, but real FRB events. We also separately generated 7500 data segments in which we simulated a pure white-noise background and injected FRBs. These training data sets were used to encourage the model to learn the correct FRB patterns. These two sets of data segments form the positive training data set while the remaining 32 500 data segments form the negative training data set.

<sup>2</sup> In the future we will re-train our model using a more physical parametrisation of the FRBs, including dispersion smearing, scattering, structures within the burst profile, and the observed frequency dependence to the burst intensity. We note that our current bandwidth is relatively small (256 MHz) and the channel bandwidths are relatively large (3 MHz). Scattering is usually small for FRB events and so the predominant effect is dispersion smearing (see e.g. Cordes & Chatterjee 2019). However, dispersion smearing can be mitigated in searches for repeating events from known FRB sources as they can be carried out using coherently de-dispersed data streams.

## 2.2. Deep neural networks architecture

A detailed introduction of deep learning and related terminology can be found in Goodfellow et al. (2016). In the following, we make use of the following terms and concepts:

- An image-based deep neural network (DNN) classifier,  $\mathcal{F}(x; \theta)$ , is a function that maps input image segments into a category,  $\hat{y} \in \{1, -1\}$ , which indicates whether the segment does or does not contain a signal of interest, respectively.
- $\mathcal{F}$  is a composite function,  $\mathcal{F}(x; \theta) = f^l(f^{l-1}(\dots(f^2(f^1(x; \theta_1); \theta_2))\dots, \theta_{l-1}); \theta_l)$ , which contains multiple internal functions,  $f^l$ . The function  $f^l$  is known as the  $l$ th “hidden layer” of the network.
- Image segments are often enhanced prior to being passed into a DNN classifier. Methods such as applying a Gaussian filter are used to smooth the input images.
- It is important to determine how well a set of parameters models the given data. This is measured using a loss function,  $\mathcal{L}(\mathcal{F}(x; \theta), y)$ , which measures the performance of the function  $\mathcal{F}(x; \theta)$ , in which  $y$  is the true label of  $x \in X$ .
- The DNN procedure obtains optimal values of the parameters  $\theta$ . An iterative method (stochastic gradient descent algorithm) is used to minimise the loss function. This relies on a step size known as the learning rate.
- The composite function,  $\mathcal{F}(x; \theta)$ , contains various hidden layers (described above) including convolutional layers, max pooling layers and fully connected layers. A convolutional layer can be thought of as a smoothing operation that is applied to the input using a matrix often referred as kernel or filter. The properties of the matrices are defined for specific features in the input images. Pooling layers reduce the dimensions of the data and hence simplify the computational complexity. Fully connected layers are directly connected to the inputs of the next layer.
- Finally, the algorithm needs to convert the numerical values of the last layer to probabilities on the various possible classifications. A softmax function is often used for this.

Our specific DNN is trained to identify single pulse events.  $\mathcal{F}(x; \theta)$ , is a function that maps our input image segment,  $\{0, 1\}^{96 \times 512}$ , into a category  $\hat{y} \in \{1, -1\}$ , which indicates that the segment does or does not contain an astronomical burst event respectively. We use the stochastic gradient descent algorithm to optimise  $\theta$  and fit the training data with a learning rate of 0.02.

To enhance the visual patterns within an image, each data segment is pre-processed using a Gaussian filter that smooths the input data, and then the segment is fed into the network for training and prediction. We have found that applying this pre-processing step improves the model training speed.

The first hidden layer in our  $\mathcal{F}$  contains two parallel convolutional blocks. The first has a  $1 \times 1$  kernel with 8 filters and the other uses a  $9 \times 9$  kernel with 32 filters. We use ReLU as the activation function for both blocks. The  $1 \times 1$  kernel is introduced to add more non-linearity to the model in order to capture patterns of various forms of RFI in image segments. The  $9 \times 9$  kernel is mainly introduced to capture the continuous patterns of the astronomical events and other non-astronomical signals. We apply a maximum pooling layer to the output of each convolutional block with a  $2 \times 2$  patch size. The output of the two maximum pooling layers are concatenated and fed into the second convolutional layer with a kernel size of  $9 \times 9$  and 128 filters. We then apply another maximum pooling layer to the output of the second convolutional layer. The network stacks two more convolutional layers with a kernel size of  $9 \times 9$  (with filter numbers of 256 and 512 respectively) and maximum pooling layers



**Table 1.** Data files processed to develop and demonstrate our algorithm.

Project	Observing semester	Sampling time (s)	File count	Integration time per file (s)	Known signals	Reference	Usage
P269	Jan 2001	0.001	2251	8400	Lorimer burst (FRB010724), and FRB010312	Kaspi et al. (2016a)	Training and testing
P269	Oct 2000	0.001	951	2818	–	Kaspi et al. (2016b)	Training
P268	Aug 1997	0.00025	4615	2100	Single pulses from known pulsars	Lyne et al. (2012a)	Training and testing
P268	May 2001	0.00025	1820	2100	FRB010621	Lyne et al. (2012b)	Testing

**Notes.** The testing data files are independent from the training data files.

before passing the output to a fully connected layer with 512 neurons. With a large number of parameters, DNN often overfits a training data set (in particular with relatively few input examples). We added a dropout layer to improve the generalisation of  $\mathcal{F}$ . An additional fully connected layer with eight neurons is stacked before a softmax function is applied to obtain the probability distribution among the event and non-event categories. In order to improve the generalization of the model for different input data collections, we use L2 regularisation. This regulariser makes the model avoid learning trivial features that only present in the training data. The DNN classifier is implemented using Tensorflow.

We trained our neural network using the real and simulated data sets that were described in the previous section. We then applied the trained model to data sets containing known events (such as the Lorimer burst and known pulsars) for demonstration and testing purposes.

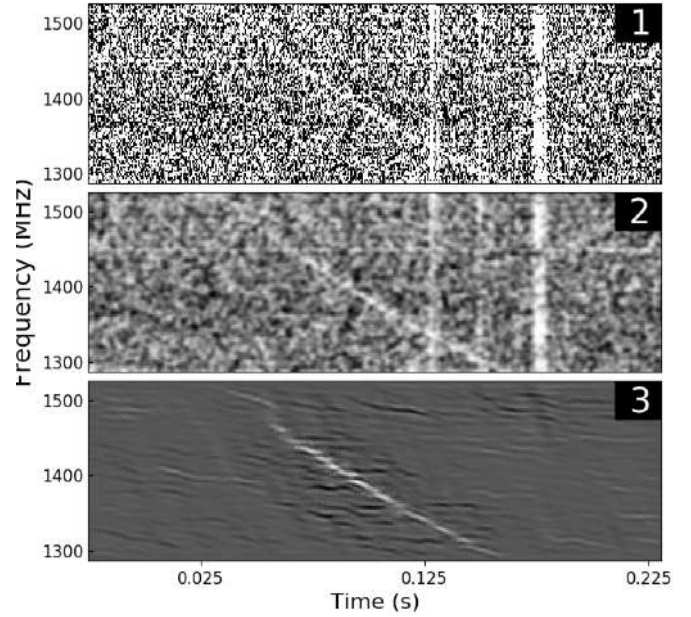
### 3. Saliency-map analysis

Saliency maps rank the pixels in the input image based on their influence on a probability score in a prediction (Simonyan et al. 2013). For deep neural networks, the influence can be calculated through the derivative of the score with respect to the input at the given pixel. To capture the variation of brightness of smoothed pixels in a pulse, we use “integrated gradients” (Sundararajan et al. 2017) to distinguish the astronomical burst signature from background noise. We consider an input image,  $x$ , that is formed by taking  $n$  steps to add a value in each pixel from a black image  $x'$  (each pixel has a value of 0). The integrated gradient of pixel  $x_i$ , denoted by  $IG_i(x)$ , is defined as:

$$IG_i(x) = (x_i - x'_i) \int_{\alpha=0}^1 \frac{\partial \mathcal{F}(x' + \alpha(x - x'))}{\partial x_i} d\alpha \quad (1)$$

where  $\mathcal{F}(x; \theta)$  is the DNN classifier and  $\alpha$  denotes the step taken on the path of changing from  $x'$  to  $x$ . The integrated gradients are able to determine how different pixels in the input image contribute to a prediction.

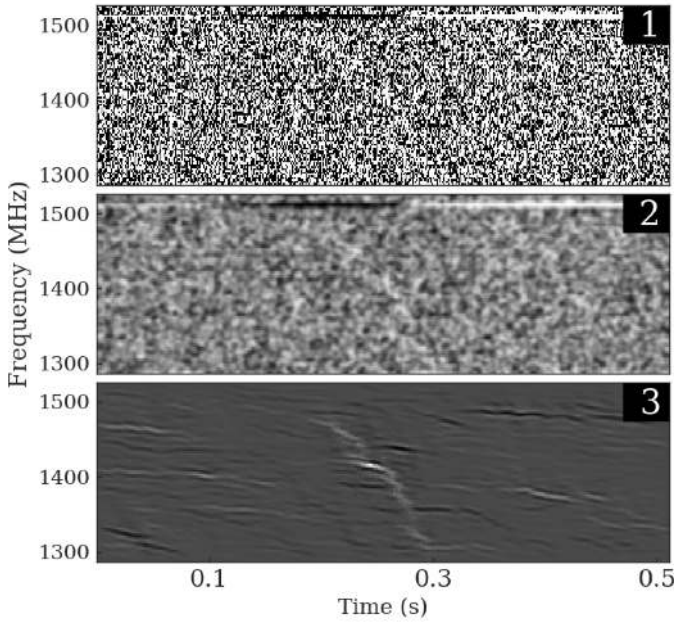
To demonstrate this process, we provide an example in Fig. 2. For this observation the telescope was pointing towards the Vela pulsar, PSR J0835–4510, and individual pulses from the pulsar are easily detectable (the figure only shows a single pulse). The top panel (labelled 1) is a segment of raw frequency-time data and clearly shows the pulse as well as three wide-band, impulsive RFI events. We show the image after smoothing with a Gaussian filter in the middle panel (labelled 2). Our machine learning classifier identified this region as containing an astrophysical event. However, we need to ensure that it has correctly



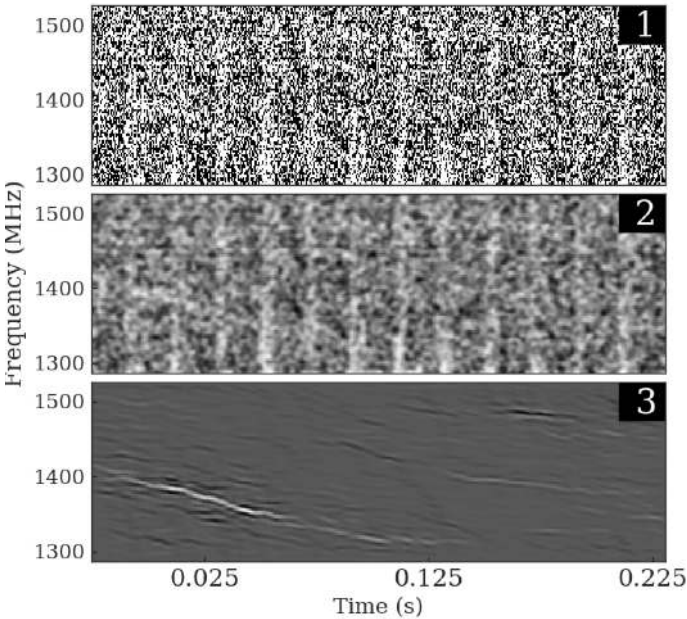
**Fig. 2.** Demonstration of the use of saliency map to identify an individual pulse from the Vela pulsar (PSR J0835–4510). *Upper panel:* raw frequency-time image. The curved feature is a single pulse from the pulsar. The vertical stripes are radio interference. *Central panel:* smoothed version of the raw data. The saliency map is shown in the *bottom panel*.

identified the pulse and not the RFI. The corresponding saliency map is shown in the bottom panel (labelled 3). The brighter a pixel is in this panel, the more important it was when predicting that the data segment contains an FRB event. The classifier has correctly identified the single pulse as a feature for its positive classification, whilst ignoring the RFI features.

Saliency maps can also be used for feature enhancement (Simonyan et al. 2013) and to understand why a particular image was not classified as an astrophysical burst event. For instance, the image in Fig. 3 was characterised as not containing an astrophysical burst event by the classifier. Panel 1 clearly shows narrow-band RFI around 1500 MHz as well as a weak signal occurring between time  $\sim 0.2$  and  $0.3$  s. This feature is not significantly enhanced in the raw data simply by smoothing the image (panel 2). As we have defined the saliency map, bright pixels correspond to regions in the image that support the hypothesis that an FRB event is present. The saliency map shows us that there is an FRB-like event present, but there is not sufficient evidence for the classifier to determine that this event is real. This highlights the possibility of using saliency



**Fig. 3.** Demonstration of the use of saliency map to determine why a plausible event was not identified as an astrophysical source. *Upper panel:* burst event in the raw data. *Central panel:* smoothed version of the raw data. The saliency map is shown in the *lower panel*.



**Fig. 4.** Single pulse event from PSR J1536–5433 that is not centred on the image and not affected by periodic radio frequency interference.

analysis to enhance features in images that have an FRB-like form, but are in some way different from the training data set (i.e. even though the algorithm was trained on ideal FRB events, the saliency maps can highlight similar but not identical features).

In our first and second examples the events occurred in the centre of the saliency map. This is coincidental, and to demonstrate the success in saliency-map analysis when the event is not directly in the centre of the image, in Fig. 4 we show an example where the procedure clearly identifies and enhances a single pulse event from PSR J1536–5433. We note that no RFI

rejection was carried out and yet the impulsive, broadband RFI clearly present in the raw data is not present in the saliency map.

In order to explore the use of saliency maps further, we have analysed a sub-set of our real and simulated data files using both a traditional PRESTO single pulse search pipeline (Ransom 2001) and our ML algorithm. The PRESTO pipeline has been described in detail by Zhang et al. (2018b, 2019). In brief, it uses `rfind` to mask strong narrow-band and short-duration broadband RFI. The `-noclip` option is turned on to avoid deleting potential bursts. `DDplan` is then used to determine the DMs to be investigated in the de-dispersion phase (set here to have 440 trials between 0 and  $5000 \text{ pc cm}^{-3}$ ). De-dispersion is carried out using `prepdata` and RFI removed using the masks produced by `rfind`.

A direct comparison of candidate lists is non-trivial and an in-depth comparison between methods and candidates will be presented elsewhere. One challenge in comparing candidate lists is that the PRESTO pipeline often produces multiple candidates for the same event (with slightly different DMs, event times, and/or widths), another challenge is that the PRESTO pipeline is extremely versatile and can be “tuned” using different input parameters. Our PRESTO-based pipeline grouped all the candidates that occurred close together in time (within a 10 ms time window). If the candidate with the highest S/N in a group has a S/N higher than seven then it was manually inspected. We found that there was an exact match between candidate lists for candidates with  $S/N > 12$ . These particular candidates were single pulses from known pulsars and the agreement with PRESTO shows that the ML algorithm is sufficient for the tests we describe here. We note that our ML algorithm is significantly faster than the PRESTO search as our process contains no de-dispersion, nor RFI mitigation stages.

The RFI-free and feature-enhanced saliency images can be used to enable a direct fit for the DM of the event and to get an estimate of the significance of the event in a way that is not affected by RFI (and without any de-dispersion steps). To show this we have determined the S/N for four examples shown in this paper as determined using the PRESTO pipeline (Zhang et al. 2018b, 2019) and compared the results with those obtained from a fit to the event in the saliency map. The raw data containing PSR J0835–4510 and PSR J1536–5433 contain impulsive broadband RFI whereas the Lorimer burst, PSR J1057–5226 and PSR J1744–3130 data sets have little detectable RFI. The S/N values are listed in Table 2, where  $S/N_1$  is the signal-to-noise ratio determined from the PRESTO pipeline and  $S/N_2$  is that determined from our pipeline. Comparison is non-trivial as the noise is well defined for the traditional analysis, but the noise in a saliency map is non-Gaussian. However, we can clearly see that the  $S/N_1$  values range from 4 (for a data set affected by RFI) to 35 (for a bright event in a clean data set), whereas all the  $S/N_2$  values are similar.

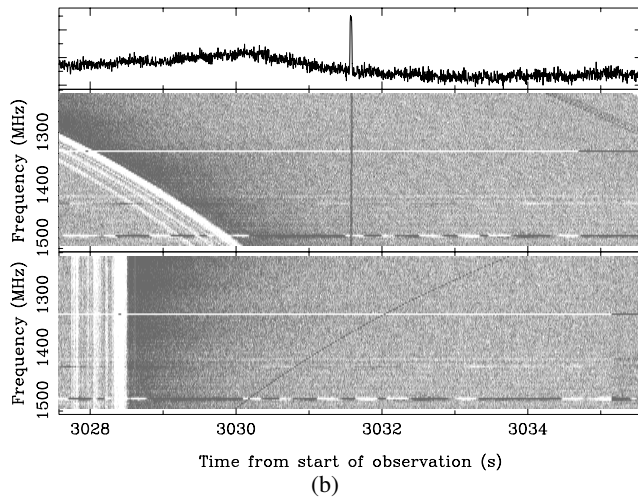
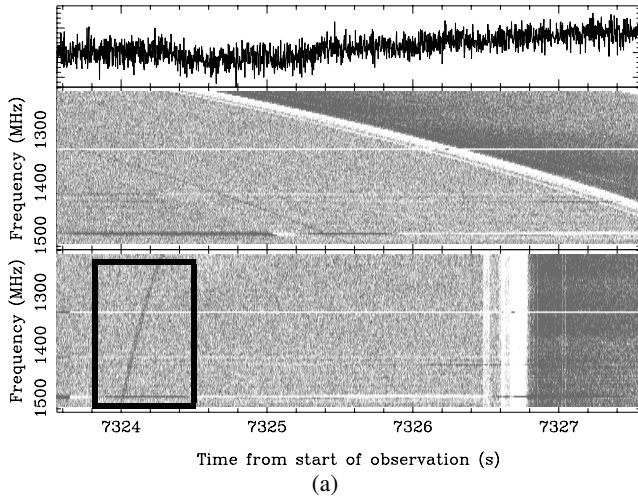
The ML algorithm produced fewer candidates than PRESTO for events with  $S/N_1$  values  $< 12$ . A future in-depth analysis will determine whether this is because the ML algorithm is missing true candidates or whether PRESTO is presenting false positives. We also found that the PRESTO-based pipeline did not identify a few of our simulated FRB events that were injected into actual data sets (i.e. they are false negative examples). These events were successfully detected using the ML method. In Fig. 5, we show two such false negative examples. We note that the panels in these images differ from the saliency-map demonstrations. Here the top panel shows the de-dispersed time series (using the DM of the PRESTO candidate), the central panel shows the dedispersed data as a function of time and frequency, and



**Table 2.** Comparison between determining S/N using the raw data file and the saliency map.

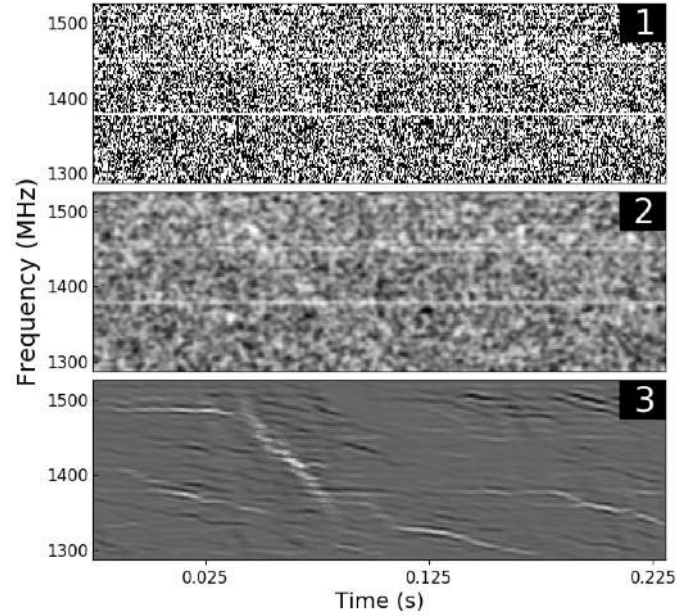
Event	$S/N_1$	nbin <sub>1</sub>	$S/N_2$	nbin <sub>2</sub>	DM ( $\text{cm}^{-3} \text{ pc}$ )
Lorimer burst	35.18	26	55.06	14	375
PSR J0835–4510	10.54	10	58.69	14	67.97
PSR J1536–5433	4.27	12	67.16	18	147.5
PSR J1057–5226	13.08	6	49.46	14	29.69
PSR J1744–3130	6.86	18	37.15	20	192.90

**Notes.** We use the definition of  $\sigma = \sum(\text{signal} - \text{noise})/\text{rms}/\sqrt{\text{nbin}}$  for our S/N value, where nbin is the best filter width in units of samples.



**Fig. 5.** False negative examples from the PRESTO single pulse search pipeline on data sets in which fake FRB signals have been injected. The sub-panels are described in the text. (a) The candidate detected by the PRESTO pipeline has a DM of  $3962 \text{ pc cm}^{-3}$ , whereas the actual signal has a DM of  $300 \text{ pc cm}^{-3}$ . (b) The S/N of a signal is distorted by RFI, which leads to a low S/N value for the FRB event.

the bottom panels shows the raw data. In Fig. 5a, the candidate produced by the PRESTO pipeline has a DM of  $3962 \text{ pc cm}^{-3}$ , whereas the simulated signal (highlighted with a black rectangle) has a DM of  $300 \text{ pc cm}^{-3}$ . Our choice of parameters when running the PRESTO pipeline failed to detect (and remove) the RFI that shadows the real signal (the S/N of the RFI is 242.19).



**Fig. 6.** Unexplained transient signal with a dispersion measure of  $41 \text{ pc cm}^{-3}$ . The raw data are shown in the *top two panels*. The saliency-map image is presented in the *bottom panel*.

Our machine learning algorithm correctly identified this simulated FRB event with a S/N of 25.9. In Fig. 5b, the measurement of the S/N of the simulated signal is affected by RFI. This candidate was found by PRESTO with a low S/N value and was filtered out by our S/N cutoff. The machine learning algorithm correctly detects this signal with a S/N of 29.0.

During the testing of our classifier we found a new single pulse from an unknown source from the observation file PM0143\_012D1.sf in the data collection P268–2001MAY. In Fig. 6, we present the raw data and the corresponding saliency map for this potential discovery, which has a right ascension and declination of 19:14:43 and +02:26:13, respectively. Panel 1 contains the raw time-frequency data, which is smoothed in panel 2. The saliency map is given in panel 3. The DM corresponding to this candidate is  $41 \text{ pc cm}^{-3}$ , which is significantly smaller than the Galactic contribution in the source direction. The traditional pulses search pipeline (such as PRESTO; Ransom 2001) identified this event with a S/N of  $\sim 7$ , but the S/N in the saliency map is 23.3 and we see no other comparable unknown event in our processing. If real (and further observations are planned of this sky region), then the source will be from a currently unknown pulsar or a rotating radio transient (RRAT; McLaughlin et al. 2006).

## 4. Conclusions

The primary goal of this paper was to highlight that saliency-map analysis can be used to provide confidence that a machine learning algorithm has identified a real astronomical event in a given data set. Producing saliency maps is not computationally intensive. Our initial implementation takes only  $\sim 1 \text{ s}$  to form a saliency map for a given candidate (and this can be further optimised). For this demonstration we have made use of archival data that have been 1 bit digitised. The results presented here are general and are directly applicable to higher-bit data streams.

We note that saliency maps are not unique to high time resolution data sets. Source finding is being carried out in

interferometric images to search for unusual source lobes and jets (Norris et al. 2019). Saliency-map analysis can be used to enhance such features in such complex images.

In summary, we have developed a new machine learning procedure for identifying astrophysical burst events in time-domain data streams. A detailed analysis of our algorithm and how it compares with more traditional search method will be published elsewhere. In this paper we have explored the use of saliency maps to identify the signatures within the data stream that contain the burst event. We have shown that the saliency maps are robust in the presence of RFI and provide a method to enhance burst-like signatures in a given data stream.

With the advent of new telescopes and improved instrumentation, the ability to detect burst events using traditional methods will become harder. The enormous data volumes from the Five-hundred-meter Aperture Spherical radio Telescope (FAST, Mickaliger et al. 2012; Li et al. 2018) and the Square Kilometre Array (SKA), among others, will require that computationally efficient algorithms be developed, and it will not be possible to view every candidate by eye. Machine learning algorithms, such as the one described here, clearly have a role to play in extracting the astrophysical information from such large data sets.

*Acknowledgements.* This work is supported by the National Natural Science Foundation of China (Grant No. 11988101, U1731238, 11725313, 11690024, 11743002, 11873067, U1731218, 11565010, 11603046, U1531246, 11703047, 11590783) and the CAS International Partnership Program No.114A11KYSB20160008 and the CAS “Light of West China” Program. The Parkes radio telescope is part of the Australia Telescope National Facility which is funded by the Australian Government for operation as a National Facility managed by CSIRO. This paper includes archived data obtained through the CSIRO Data Access Portal (<http://data.csiro.au>). This work was supported by a China Scholarship Council (CSC) Joint PhD Training Program grant and the National Natural Science Foundation of China. This project was supported by resources and expertise provided by CSIRO IMT Scientific Computing.

## References

- Barsdell, B. R., Bailes, M., Barnes, D. G., & Fluke, C. J. 2012, *MNRAS*, **422**, 379
- CHIME/FRB Collaboration (Amiri, M., et al.) 2019a, *Nature*, **566**, 235
- CHIME/FRB Collaboration (Andersen, B. C., et al.) 2019b, *ApJ*, **885**, L24
- Connor, L., & van Leeuwen, J. 2018, *AJ*, **156**, 256
- Cordes, J. M., & Chatterjee, S. 2019, *ARA&A*, **57**, 417
- Devine, T. R., Goseva-Popstojanova, K., & McLaughlin, M. 2016, *MNRAS*, **459**, 1519
- Farah, W., Flynn, C., Bailes, M., et al. 2019, *MNRAS*, **488**, 2989
- Goodfellow, I., Bengio, Y., & Courville, A. 2016, *Deep Learning* (MIT Press)
- Hobbs, G., Miller, D., Manchester, R. N., et al. 2011, *PASA*, **28**, 202
- Hotan, A. W., van Straten, W., & Manchester, R. N. 2004, *PASA*, **21**, 302
- Kaspi, V., Manchester, R., & Lyne, A. 2016a, *Parkes Observations for Project P269 Semester 2001JANT*
- Kaspi, V., Manchester, R., & Lyne, A. 2016b, *Parkes Observations for Project P269 Semester 2000OCTT*
- Li, D., Wang, P., Qian, L., et al. 2018, *IEEE Microw. Mag.*, **19**, 112
- Lorimer, D. R., Bailes, M., McLaughlin, M. A., Narkevic, D. J., & Crawford, F. 2007, *Science*, **318**, 777
- Lyne, A., Manchester, R., & Camilo, F. 2012a, *Parkes Observations for Project P268 Semester 1997AUGT*
- Lyne, A., Kramer, M., & Manchester, R. 2012b, *Parkes Observations for Project P268 Semester 2001MAYT*
- Manchester, R. N., Fan, G., Lyne, A. G., Kaspi, V. M., & Crawford, F. 2006, *ApJ*, **649**, 235
- McLaughlin, M. A., Lyne, A. G., Lorimer, D. R., et al. 2006, *Nature*, **439**, 817
- Men, Y. P., Luo, R., Chen, M. Z., et al. 2019, *MNRAS*, **488**, 3957
- Michilli, D., Hessels, J. W. T., Lyon, R. J., et al. 2018, *MNRAS*, **480**, 3457
- Mickaliger, M. B., Lorimer, D. R., Boyles, J., et al. 2012, *ApJ*, **759**, 127
- Montavon, G., Samek, W., & Muller, K.-R. 2018, *Digit. Signal Process.*, **73**, 1
- Norris, R. P., Salvato, M., Longo, G., et al. 2019, *PASP*, **131**, 108004
- Pan, Z., Hobbs, G., Li, D., et al. 2016, *MNRAS*, **459**, L26
- Petroff, E., Keane, E. F., Barr, E. D., et al. 2015, *MNRAS*, **451**, 3933
- Platts, E., Weltman, A., Walters, A., et al. 2019, *Phys. Rep.*, **821**, 1
- Ransom, S. M. 2001, PhD Thesis, Harvard University, USA
- Simonyan, K., Vedaldi, A., & Zisserman, A. 2013, *ArXiv e-prints* [arXiv:1312.6034]
- Spitler, L. G., Scholz, P., Hessels, J. W. T., et al. 2016, *Nature*, **531**, 202
- Sundararajan, M., Taly, A., & Yan, Q. 2017, *Proceedings of the 34th International Conference on Machine Learning – Volume 70, ICML’17 (JMLR.org)*, 3319
- Zhang, Y. G., Gajjar, V., Foster, G., et al. 2018a, *ApJ*, **866**, 149
- Zhang, S.-B., Dai, S., Hobbs, G., et al. 2018b, *MNRAS*, **479**, 1836
- Zhang, S.-B., Hobbs, G., Dai, S., et al. 2019, *MNRAS*, **484**, L147