



UvA-DARE (Digital Academic Repository)

Learning to detect radio frequency interference in radio astronomy without seeing it

Mesarcik, M.; Boonstra, A.-J.; Rangelova, E.; van Nieuwpoort, R.V.

DOI

[10.1093/mnras/stac2503](https://doi.org/10.1093/mnras/stac2503)

Publication date

2022

Document Version

Final published version

Published in

Monthly Notices of the Royal Astronomical Society

License

CC BY

[Link to publication](#)

Citation for published version (APA):

Mesarcik, M., Boonstra, A.-J., Rangelova, E., & van Nieuwpoort, R. V. (2022). Learning to detect radio frequency interference in radio astronomy without seeing it. *Monthly Notices of the Royal Astronomical Society*, 516(4), 5367-5378. <https://doi.org/10.1093/mnras/stac2503>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)

Learning to detect radio frequency interference in radio astronomy without seeing it

Michael Mesarcik,¹★ Albert-Jan Boonstra,² Elena Ranguelova³ and Rob V. van Nieuwpoort^{1,3}

¹*Informatics Institute, University of Amsterdam, Science Park 904, NL-1098 XH Amsterdam, the Netherlands*

²*ASTRON, the Netherlands Institute for Radio Astronomy, Oude Hoogeveensedijk 4, NL-7991 PD Dwingeloo, the Netherlands*

³*eScience Center, Science Park 140, NL-1098 XG Amsterdam, the Netherlands*

Accepted 2022 September 1. Received 2022 August 18; in original form 2022 July 1

ABSTRACT

Radio frequency interference (RFI) corrupts astronomical measurements, thus affecting the performance of radio telescopes. To address this problem, supervised-segmentation models have been proposed as candidate solutions to RFI detection. However, the unavailability of large labelled data sets, due to the prohibitive cost of annotating, makes these solutions unusable. To solve these shortcomings, we focus on the inverse problem: training models on only uncontaminated emissions, thereby learning to discriminate RFI from all known astronomical signals and system noise. We use nearest latent neighbours – an algorithm that utilizes both the reconstructions and latent distances to the nearest neighbours in the latent space of generative autoencoding models for novelty detection. The uncontaminated regions are selected using *weak labels* in the form of RFI flags (generated by classical RFI flagging methods) available from most radio astronomical data archives at no additional cost. We evaluate performance on two independent data sets, one simulated from the Hydrogen Epoch of Reionization Array (HERA) telescope and the other consisting of real observations from the Low-Frequency Array (LOFAR) telescope. Additionally, we provide a small expert-labelled LOFAR data set (i.e. strong labels) for evaluation of our and other methods. Performance is measured using the area under the receiver operating characteristic (AUROC), area under precision–recall curve (AUPRC), and the maximum F1-score for a fixed threshold. For the simulated HERA data set, we outperform the current state of the art across all metrics. For the LOFAR data set, our algorithm offers both a 4 per cent increase in AUROC and AUPRC at the cost of increasing the false negative rate, but without any manual labelling.

Key words: methods: data analysis – techniques: interferometric.

1 INTRODUCTION

Radio frequency interference (RFI) is a growing concern for radio astronomy due to the proliferation of electronic equipment that depends on electromagnetic emissions. Radio frequency radiation from astronomical sources is extremely faint relative to emissions from man-made systems such as radars, telecommunication devices, large satellite constellations (Hainaut & Williams 2020), and more. Despite international regulation to ensure radio-quiet zones and limit transmission power of emitters, there are still concerns about RFI hampering radio astronomy.

For this reason, approaches for RFI detection and mitigation have become a necessity in modern radio observatories. RFI pipelines are commonly deployed at telescopes performing RFI detection and mitigation in a post-correlation setting. Traditionally, algorithms such as CUMSUM (Baan, Fridman & Millenaar 2004), singular value decomposition (Offringa et al. 2010), wavelet-based methods (Maslakov et al. 1998), and AOfagger (Offringa, Van De Gronde & Roerdink 2012) have been used. These RFI detection algorithms are widely implemented for real-time RFI detection at observatories around the world (Kildal, Baker & Hagfors 1991; van

Haarlem et al. 2013; Sokolowski et al. 2015; Foley et al. 2016). In effect, all archived data from these instruments contain automatically generated RFI masks, which can be accessed with no additional cost.

Recent advances in machine learning have made data-driven approaches unprecedentedly suitable for RFI detection. Most machine learning approaches to RFI detection have been based on supervised learning using U-Net (Ronneberger, Fischer & Brox 2015) and its derivatives (Akeret et al. 2017b; Kerrigan et al. 2019; Sadr et al. 2020; Yang et al. 2020). Research has shown that these are promising methods, significantly outperforming classical approaches. However, in reality, supervised methods require significant amounts of expert-labelled time/frequency data, which are not available in practice due to the related cost.

As a result, recent models are trained and evaluated using simulators or flags generated by classical methods, with limited experimentation on real expert-labelled data sets. This is problematic as the effectiveness of these methods on unseen data is difficult to measure and predict. Furthermore, recent machine learning-based methods have not been well integrated into telescope pipelines as the cost of labelling is prohibitive to many instrument operators.

To solve these problems, we propose an unsupervised learning method based on the nearest-latent-neighbours (NLN) algorithm (Mesarcik et al. 2022). This approach leverages novelty detection to perform RFI detection. This is achieved using a generative

* E-mail: m.mesarcik@uva.nl

model trained on uncontaminated (RFI-free) data to detect *novel* RFI contaminated emissions. Interestingly, this formulation is effectively the inverse of how existing deep learning-based methods are trained.

In this work, RFI is detected by measuring the difference between small subregions (patches) of spectrograms that are known to *not* contain RFI and the patches being evaluated. To select RFI-free data, we break each spectrogram into a number of equally sized patches and use their associated AOFlagger-based flags to locate the instances which contain RFI. In doing so, we do not have to incur the cost of extensive labelling as the AOFlagger masks are readily available. It must be noted that NLN can tolerate high false positive rates in the training masks due to the use of the inverse problem. However, underflagging can cause undesired effects.

We show that our method outperforms existing supervised models on several benchmarks, using less data for training. Furthermore, we demonstrate that if supervised state-of-the-art methods are trained with weak labels, they typically overfit to the training data and do not generalize to unseen examples. Our approach does not suffer from these problems.

Additionally, as the landscape of RFI emissions changes over time, we expect supervised methods to be continually retrained as future emitters occupy newer frequency bands. Conversely, the emissions from celestial bodies will remain fairly consistent over the same time-scale, effectively meaning that our method will not have to undergo retraining.

We make the following contributions in this work: (1) a novel unsupervised learning-based approach to RFI detection in radio astronomy; (2) an evaluation of the effectiveness of using AOFlagger-generated ground truth for training of machine learning-based RFI detection algorithms; and (3) an expert-labelled data set that can be used for comparison and development of novel RFI detection algorithms.

This paper begins with the critical discussion of the existing literature concerning RFI detection in radio astronomy in Section 2. Section 3 explains how the NLN algorithm was adapted to work for RFI detection. In Section 4, we explain our data selection strategy and outline the expert-labelled data set used for evaluation of this work. Finally, we present our results and conclusions in Sections 5 and 6.

2 RELATED WORK

Machine learning for RFI detection is an actively researched field. Numerous works offer a variety of radio astronomy-specific modifications to improve accuracy of detection. Furthermore, machine learning-based anomaly and novelty detection have been applied extensively in astronomy for purposes of novel galaxy morphology detection, transient detection, exoplanet discovery, and more. However, there have been a few attempts to apply novelty detection to the RFI detection problem. In this section, we document the latest astronomical developments in both RFI detection and novelty detection.

In this work, we omit analysis of and comparison with classical techniques that rely on spatial filtering, high-order statistics, and subspace decomposition (Fridman & Baan 2001). This is because these techniques typically require significant fine tuning on specific telescopes, thus making comparison challenging.

2.1 RFI detection in the deep learning era

Semantic segmentation is at the heart of the deep learning-based RFI detection, with U-Net (Ronneberger et al. 2015) and derivatives

acting as the architectural backbone of recent research. The purpose of semantic segmentation is to determine the pixel-precise regions where a specific class exists – in this case RFI. Architecturally, U-Net is a convolutional neural network, with an encoder–decoder pair that share activations between the two stages. It is trained in a supervised manner, requiring pixel-level Boolean masks per spectrogram.

The first application of U-Net to radio astronomy-based RFI detection is reported in seminal work by Akeret et al. (2017b). The network is trained and evaluated on the magnitude of spectrograms obtained from both simulated data and real data from a signal antenna from the Bleien Observatory (Chang et al. 2017). Interestingly, the models are trained using masks obtained from a classical flagging approach. We show in Section 5 that this is not ideal, as supervised methods tend to overfit to the weak-label-based ground truth. Additionally, this work makes use of the HIDE & SEEK radio astronomical data simulator (Akeret et al. 2017a). We find the use of this simulator problematic because the ground truth needs to be determined by user-defined thresholds of the residual RFI maps as described by Sadr et al. (2020). Due to this, other works such as Yang et al. (2020) do not describe the threshold used for evaluation, making comparison extremely difficult.

To counteract the issue of overfitting to the potentially incorrect labels, we focus on the inverse problem. We train a model to represent all non-RFI signals, such that any deviation from the learnt representations is flagged as RFI. It must be noted that this approach depends on the assumption that the number of false positives of classical RFI detectors is higher than the number of false negatives. In other words, most RFI is flagged as RFI, but other features may be incorrectly flagged. A detailed analysis of this effect can be found in Section 5.2.

To address the limitations of the HIDE simulator, we use the Hydrogen Epoch of Reionization Array (HERA) simulator (DeBoer et al. 2017). It gives a more granular control over the simulated interference, enabling multiple classes to be generated with precise ground truth without specifying a threshold. Furthermore, we train our model on multiple observations obtained from the Low-Frequency Array (LOFAR) Long Term Archive (LTA; van Haarlem et al. 2013) and evaluate it on a subset of expert-labelled examples.

Kerrigan et al. (2019) document the use and evaluation of a U-Net derivative for RFI detection on both real and simulated data from the HERA telescope (DeBoer et al. 2017). The authors propose to include both the magnitude and phase as separate components of the model, for better generalization to alternative representations. It must be noted that both Mesarcik et al. (2020) and Sadr et al. (2020) show that minimal improvements are obtained through using both the magnitude and phase representations of the complex visibilities. For this reason, we only use the magnitude portion of our data for RFI detection. Additionally, we extend the evaluation done by Kerrigan et al. (2019) by considering how supervised RFI detection algorithms generalize to unseen classes of RFI, what we refer to as out-of-distribution (OOD) RFI.

Transfer learning (Tang, Scaife & Leahy 2019) or domain adaptation (Farahani et al. 2021) have been shown to be powerful tools when there are limited labelled data. Sadr et al. (2020) show that it is possible to train a U-Net inspired architecture, called *R-Net*, on simulated data and then using a small sample of expert-labelled data to adapt the model’s domain from simulated to real world data. However, we find that R-Net does not offer significant improvements over the standard U-Net, when evaluating on real data without transfer learning, as shown in Section 5. Similarly, Yang et al. (2020) propose *RFI-Net*, a residual modification of U-Net that offers better training stability when using a significantly deeper network.

Generative models have also been used for RFI detection in the context of radio astronomy. In a work by Vinsen, Foster & Dodson (2019), it has been shown that generative adversarial networks (GANs) can be used for RFI detection. However, this research is limited in its evaluation and does not offer a practical way to obtain pixel-precise predictions of RFI. Finally, Vos et al. (2019) offer a significantly different paradigm for RFI detection using GANs. Here, the authors propose a source-separation approach that uses two separate generators to distinguish astronomical signals from RFI. However, this method requires significant supervision, as the model needs access to the mixture as well as the separated RFI and astronomical sources during training. We find this requirement prohibitive, as to obtain these source separations for real data is extremely costly. We show that generative models can be used without the cost of supervision, by treating novelty detection as a downstream task.

2.2 Novelty detection in radio astronomy

Machine learning-based novelty and anomaly detection have gained significant attention in radio astronomy for a number of different use cases. Topics such as transient detection (Malanchev et al. 2021), the search for extraterrestrial intelligence (Zhang et al. 2018), and detecting outliers in radio galaxy morphology (Margalef-Bentabol et al. 2020) have been researched.

Novelty detection in machine learning is typically done in a two-stage process, where a model is first fitted to the in-lying classes and then a decision boundary is found in the learnt-representation space of the model. In a work by Villar et al. (2021), outlying extragalactic transients are detected using an isolation forest (Tony Liu, Ming Ting & Zhou 2008), whereas Lochner & Bassett (2021) propose using the local outlier factor (Breunig et al. 2000) for detecting anomalies in light curves or dynamic spectra. To the best of our knowledge, there have been no attempts at applying these techniques to RFI detection in the context of radio astronomy.

The main problem with directly applying latent-variable novelty detection models to RFI detection is the resolution of the predictions. The models used for novelty detection in radio astronomy produce a scalar output for a single data point (spectrogram, radio image, etc.). This is problematic as RFI detection algorithms must produce a Boolean mask of the pixel-precise regions of where RFI is located in a given input. Therefore, we propose to use the NLN algorithm, an approach that combines both latent measures of difference and the pixel-precise reconstruction errors from a generative autoencoding model (Mesarcik et al. 2022b). This is in contrast to Harrison et al. (2019), where particular features are extracted from each spectrogram to determine the novel differences in transmission power and frequency range. However, the method does not produce pixel-precise segmentation maps of the detected RFI.

3 METHOD

In this work, we use NLN for RFI detection. This is motivated by several factors: (1) obtaining sufficient labels for supervised segmentation of RFI has a significant overhead; (2) existing supervised techniques overfit to flags from classical methods such as AOFlogger, leading to suboptimal performance on unseen data; (3) the ever-changing landscape of RFI requires continual labelling and training efforts to enable supervised approaches to capture new temporal and spectral RFI structures.

We show that if the traditional RFI detection problem is inverted, we can effectively address the supervised RFI detection issues. All models and code used in this work are publicly available.¹

3.1 Model definition and training

For some complex visibility $V(v, \tau, b)$ and the corresponding ground-truth mask for the interference $G(v, \tau, b)$, the training objective for supervised RFI detection can be formulated as follows:

$$\mathcal{L}_{\text{sup}} = \min_{\theta_m} \mathcal{H}(m_{\theta_m}(V(v, \tau, b)), G(v, \tau, b)). \quad (1)$$

Here, m is a function with learnable parameters θ_m and \mathcal{H} is the entropy-based similarity between the model prediction and the ground truth. This problem is well posed and has been used across multiple domains to effectively train classifiers. However, it relies on learning a model of the RFI using ground-truth labels, which are in practice hard to obtain.

In this work, we train a model of everything other than the interference, so that we can perform RFI detection as a downstream task. This is done by first training a discriminative autoencoder, on $n \times n$ -sized uncontaminated regions (also known as patches) of the visibility, $V_{n \times n}(v, \tau, b)$. We select these regions using the *weak labels* generated by a classical method such as AOFlogger.

First, we define the encoder f that maps from the visibility space \mathbb{R}^2 to a latent space \mathbb{L} , such that

$$\mathbf{z} = f_{\theta_f}(V_{n \times n}(v, \tau, b)), \quad f: \mathbb{R}^2 \rightarrow \mathbb{L}. \quad (2)$$

Here, \mathbf{z} is a low-dimensional projection of an $n \times n$ patch that contains no interference and θ_f are the learnable parameters of the encoder. Furthermore, we define the decoder g that maps back from the low-dimensional projection to the visibility space, such that

$$\widehat{V}_{n \times n}(v, \tau, b) = g_{\theta_g}(\mathbf{z}), \quad g: \mathbb{L} \rightarrow \mathbb{R}^2, \quad (3)$$

where θ_g are the parameters of the decoder. We simultaneously train the encoder and decoder using the reconstruction loss,

$$\mathcal{L}_{\text{recon}} = \min_{\theta_f, \theta_g} \mathcal{H}(V_{n \times n}(v, \tau, b), \widehat{V}_{n \times n}(v, \tau, b)). \quad (4)$$

We use mean square error (MSE) to train a standard autoencoder. Typically, MSE-based reconstruction losses produce blurry outputs, which may affect the quality of the predicted RFI masks. The blurriness is a result of back-propagating the gradient from the average pixel-wise error, which prohibits the autoencoder from producing high-frequency details for given inputs. In order to counteract this problem, we define a discriminator, d , that acts as a regularizer on the decoder's output. This discriminative loss is back-propagated through the decoder, such that it learns to produce not only low-frequency details, but also the discriminative features (Srivastava et al. 2017). Furthermore, the discriminator enables the autoencoder to be used as a generative model (Larsen et al. 2016). The discriminator maps from \mathbb{R}^2 to a classification on the interval $[0, 1]$, effectively trying to determine whether the input is generated or sampled from the original data set. In this case, the original data set is the uncontaminated patches selected using the weak labels. The discriminative loss is given by

$$\begin{aligned} \mathcal{L}_{\text{disc}} = \min_{\theta_d} \mathbb{E}[\log(d_{\theta_d}(V_{n \times n}(v, \tau, b)))] \\ + \mathbb{E}[\log(1 - d_{\theta_d}(\widehat{V}_{n \times n}(v, \tau, b)))] \end{aligned} \quad (5)$$

¹<https://github.com/mesarcik/RFI-NLN>

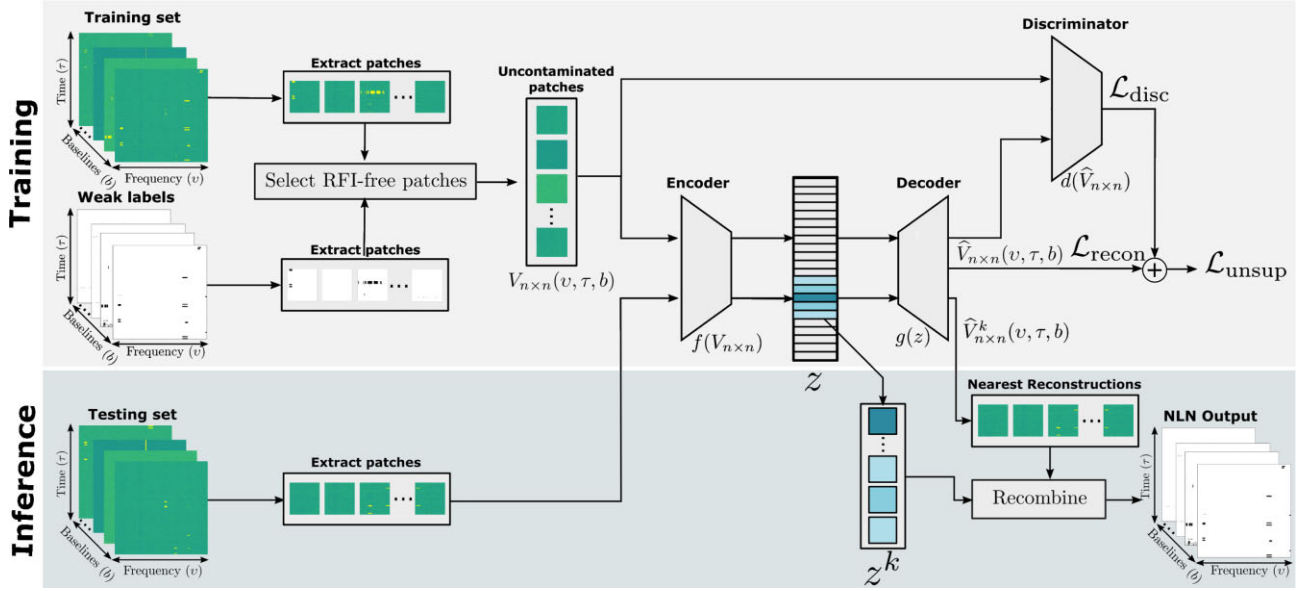


Figure 1. Block diagram of the training and inference procedures of NLN-based RFI detector. Here, we use a discriminative autoencoder as the backbone of the architecture. The top half of the figure shows the training procedure, whereas the bottom half illustrates how NLN is used for inference. The recombination of latent distances z^k and nearest reconstructions $\hat{V}^k(v, \tau, b)$ is performed according to equation (6).

We train the discriminator simultaneously with the decoder, such that the total loss is $\mathcal{L}_{\text{unsup}} = \alpha \mathcal{L}_{\text{recon}} + (1 - \alpha) \mathcal{L}_{\text{disc}}$, where α is a hyper-parameter between 0 and 1 that determines how much of an effect the discriminative and reconstruction losses have on the training, respectively. An illustration of the training procedure for the discriminative autoencoder is shown in the top half of Fig. 1.

3.2 NLN for RFI detection

Given some trained latent-variable model such as the discriminative autoencoder, $g(f(V))$, we need to formulate a measure of similarity or difference between the learnt distribution of RFI-free patches and those that are unseen. In practice, several options exist such as pixel-level difference (Akçay, Atapour-Abarghouei & Breckon 2019), structured-similarity measure (Bergmann et al. 2019b), residual measures (Schlegl et al. 2017), purely latent measures (Bergman, Cohen & Hoshen 2020), and many more. The most important factor when selecting a measure for RFI detection is the resolution of the output. For example, using a purely latent measure would result in the resolution of the output RFI masks to be fixed by the resolution of each patch, as shown in Fig. 2(e). However, using a pixel-level difference may cause the predictions to be sensitive to noise.

To counteract this problem, we propose using a distance function that utilizes both latent and pixel-wise measures of difference, namely NLN. NLN is a novelty detection technique that works by performing a nearest-neighbour lookup in the latent space of generative models. At training time, it operates as a standard discriminative autoencoder, training on the aforementioned loss. During inference, a test sample is given to the model, with the objective to determine which parts (if any) of the input sample are novel. A combination of two metrics is used: the first measures the latent distance from the given test sample to its nearest neighbours from the distribution of in-lying data, as illustrated in Fig. 2(e), and the other is the reconstruction error between the given sample and the reconstructions of all its neighbours found in the latent space. Figs 2(b) and (c) demonstrate that when the autoencoder is trained on only RFI-free data, it is capable of

only reconstructing the *non-novel* astronomical signals and cannot generate RFI found in the input. In effect, the reconstruction error shown in Fig. 2(d) has a higher dynamic range than the input. For more details and analysis of the method, see Mesarcik et al. (2022b).

We modify the original NLN-distance function such that the latent distances are used as a coarse selection for the higher resolution pixel-based error. An illustration of this selection mechanism is shown in Fig. 2(f). The modified NLN measure is the reconstruction error of a test sample's nearest neighbours multiplied by its latent-distance vector, as given by

$$D_{\text{NLN}} = \left(\frac{1}{K} \sum^K |V_{i,n \times n}(v, \tau, b) - g_{\theta_g}(z_i^k)| \right) \times D_{\text{latent}}, \quad (6)$$

where k is the nearest neighbour of the i th sample in the latent space given by z . The nearest neighbours are selected through the K -nearest-neighbours algorithm using the default implementation of FAISS (Johnson, Douze & Jegou 2021). Furthermore, D_{latent} is the thresholded mean latent-distance vector of the i th query patch and its k RFI-free neighbours, as given by

$$D_{\text{latent}} = \begin{cases} 1, & \frac{1}{K} \sum^K |z_i - z_i^k| \geq T. \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

We treat both K and T as hyper-parameters of our algorithm and determine them experimentally across our data sets. In effect, the latent-distance function offers a coarse resolution view of the RFI, and the reconstruction error offers a finer grained resolution. It must be noted that the only additional overhead of NLN is that it requires the representations of the training set to be stored. An illustration of the inference mode of the NLN algorithm is shown in the bottom half of Fig. 1.

3.3 Architectural considerations

We use a strided convolutional architecture for the encoder, decoder, and discriminator. Both the encoder and decoder have the same

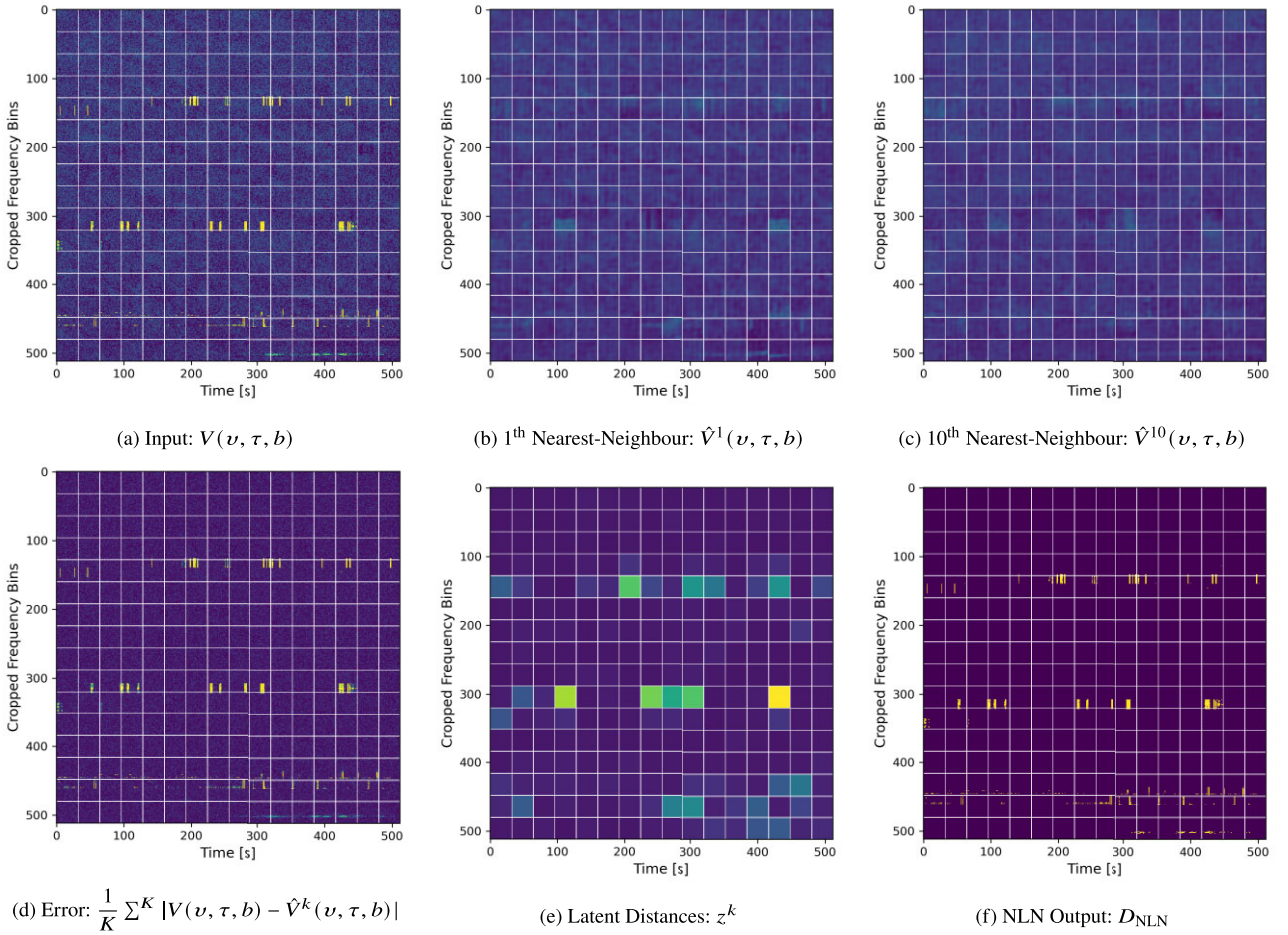


Figure 2. Stages of NLN-based RFI detection on the 30th sample of the LOFAR data set. The white grid illustrates the re-composition of the 32×32 patches to their corresponding locations in the original spectrogram. Each subplot reflects a part of the modified NLN algorithm from equation (6).

architecture except that the decoder uses transposed convolutions in place of the encoder’s convolutional layers. Furthermore, the discriminator uses the same architecture as the encoder, except for the final layer, which is a linear layer with a sigmoid activation for the discriminator.

Several parameters of the architecture are constrained by the chosen patch size and stride width. We find that a patch size of 32×32 generally exhibits the best performance, as shown in Fig. 8; this limits both the depth and latent dimensionality of our networks. For this reason, the three networks have two convolutional layers with 3×3 filters and a stride of 2. Each convolutional layer is followed by a batch normalization layer and a dropout layer with a rate of 5 per cent to regularize the network. Lastly, the convolutional output is projected to a specified latent dimensionality by a linear layer.

A base number of filters of 32 is used for the AE and is increased or decreased on each subsequent layer by a factor of 2. We use *ReLU* activations for all models and they are trained for 100 epochs using ADAM (Kingma & Ba 2015) with a learning rate of 1×10^{-4} .

4 DATA SELECTION AND PRE-PROCESSING

Existing machine learning-based approaches rely on significant amounts of labelled data for training and evaluating the models. By inverting the RFI detection problem, we do not need explicit training

Table 1. Attributes of each data set used for training and evaluation. The low test-train-ratio is due to the use of weak labels generated from classical methods, which are not directly used in the NLN training.

Data set	No. of baselines	No. of training samples	No. of test samples	% RFI
HERA	28	420	140	2.76
LOFAR	2775	7500	109	1.26

labels, but rather rely on the weak labels that typically come without additional cost from data archives such as the LOFAR LTA (van Haarlem et al. 2013). This means that we only need very few expert-labelled examples for evaluation of our model while training on a large data set as shown in Table 1.

We use two different data sets from two different telescopes to evaluate our work: simulated data from HERA (DeBoer et al. 2017) and calibration data from the LOFAR LTA. We use AOFlogger-based weak labels for training all models shown in this work on both the HERA and LOFAR data sets. For evaluation with HERA, we use the ground truth supplied by the simulator, whereas for the LOFAR we hand-annotate a selection of baselines obtained from the archive. We have made all data sets used in this work publicly available.²

²<https://zenodo.org/record/6724065>

4.1 Simulated HERA data set

The HERA simulator³ generates complex spectrograms from a simulated radio telescope. It uses models of diffuse sources, multiple types of RFI emissions, and systematic models of the HERA telescope for parameters such as antenna cross-coupling, bandpass effects, and more. Importantly for this work, the simulator gives operators a fine-grained control over the generated RFI types as well as their pixel-precise ground-truth maps. Thanks to these properties, we can use the HERA data for the validation of our approach.

In this work, we simulate a hexagonal array with 14.6 m between each station, as specified by the 19-element array (DeBoer et al. 2017). To synthesize our data set, we create a 30 min observation with an integration time of 3.52 s and a bandwidth of 90 MHz from 105 to 195 MHz (with 512 frequency channels). The specific integration interval is used to ensure that the resulting spectrograms are square; this is done to simplify the arithmetic of creating and reconstructing the spectrograms from patches. Furthermore, we use the default number of diffuse galactic emissions specified in Kerrigan et al. (2019) for the H1C observation season from 2017 to 2018 and we include additive noise at a temperature of 180 K in our simulations.

In the generation of our training set, we individually synthesize multiple RFI emissions based on the models specified in Kerrigan et al. (2019). These being narrow-band continuous emissions modelled satellite communication such as *ORBCOM*, broad-band transient emissions that imitate events such as lightning, as well as narrow-band burst RFI based on ground communication. Additionally, we include random single time–frequency *blips*.

Using the hexagonal array, we simulate 28 baselines of both auto and cross-correlations. We repeat this 20 times to obtain 560 complex spectrograms to train and evaluate our models with. This approach was deemed more appropriate than simulating a single long observation as performed in Kerrigan et al. (2019), as there is more diversity in the RFI landscape from multiple initializations of the simulator. The simulated data have an RFI occupancy rate of 2.76 per cent and is split into 75 per cent for training and 25 per cent for testing, as reported in Table 1.

The simulated data are pre-processed before training and evaluation. For simplicity purposes, we only use the magnitude of the complex visibilities. To deal with the high dynamic range, we clip the data in the range $[|\mu - \sigma|, \mu + 4\sigma]$ and take the natural log of the clipped spectrograms. Finally, before training, we standardize our data between 0 and 1 to ensure that the gradients while training do not explode or vanish.

4.2 LOFAR

As previously mentioned, we use publicly available data from the LOFAR LTA (van Haarlem et al. 2013). Five measurements were randomly selected from 2017 to 2018 for the evaluation of our model. We select calibration measurements that point at strong radio sources, using 51 stations in the band 120–190 MHz for 600 s. The precise details of the observations are available in Mesarcik et al. (2022).

A common challenge in applying machine learning to radio astronomy is the amount of data generated by observations. This is especially problematic when training models with limited GPU memory. For example, the five raw measurement sets used in this work are 1.7 TB cumulatively in size. Therefore, in order to decrease

the data size, we use only the magnitude of the first Stokes parameter and randomly sample 1500 baselines of each observation. This effectively reduces the data set size to ~ 10 GB.

For purposes of further reducing the training memory footprint of the data and simplifying the arithmetic of reconstructing the spectrograms from their respective patches, we first create approximately square spectrograms. As there are 599 time samples per observation and 16 subbands per channel, we concatenate each 44 consecutive subbands together to form spectrograms of size 599×616 . Additionally, we discard the first and last channels of each subband due to bandpass effects. Finally, we randomly crop the resulting spectrograms to 512×512 . This is done, as cropping gives an equal representation of all frequency bands.

For the evaluation of the models on the LOFAR data set, we randomly select 109 baselines for expert labelling. This is in line with the number of baselines used for evaluation in the simulator-based setting. As noted in Kerrigan et al. (2019), and Sadr et al. (2020), there are often discrepancies between the AOFlagger masks and those given by an expert, and this is highlighted in Fig. 3. For validation and evaluation of this work, we treat the hand-annotated examples as the ground truth, as described in Section 5. Furthermore, we use the AOFlagger masks associated with the measurement sets from the LTA for training our models.

In Table 1, we report the percentage of RFI contamination and the data set sizes. We ensured that the contamination is in line with what is reported in van Haarlem et al. (2013). Furthermore, as our method does not rely on human-labelled examples to train, we only require a small number of expert-labelled examples to evaluate the performance of our models.

We pre-process the LOFAR data in a similar manner to the simulated data. We first clip the data in the range $[|\mu - \sigma|, \mu + 20\sigma]$ and followed by the natural logarithm. We finally standardize the data between 0 and 1 to ensure that the gradients while training do not explode or vanish.

5 RESULTS

We evaluate the performance of NLN applied to RFI detection experimentally on both simulated and real data from HERA and LOFAR, respectively. Unlike previous works, we perform our evaluation in a two-step process, first by finding an appropriate AOFlagger threshold to generate the training annotations and then calculating the respective performance of each model on the real ground truth. To ensure the correctness of the evaluation, we use the ground-truth masks from the simulator in the test set for the HERA models and human-generated annotations for the testing of the LOFAR models.

5.1 Evaluation methodology

Following previous efforts to evaluate the performance of RFI detection (Offringa et al. 2012; Akeret et al. 2017b; Kerrigan et al. 2019; Sadr et al. 2020), we use the three most common metrics: the area under the receiver operating characteristic (AUROC) score, area under precision–recall curve (AUPRC), and F1-score. The AUROC metric evaluates the ratio of true positive rate (TPR) and false positive rate (FPR) across several thresholds. In this case, the TPR is the fraction of RFI that is correctly classified as RFI and the FPR is the fraction of misclassified signals. The AUPRC metric gives the ratio of precision and recall when the model’s output is evaluated across several thresholds. In this case, precision refers to the fraction of correctly classified RFI across all RFI predictions, and the recall is simply the TPR. Finally, the F1-score is the harmonic mean of

³https://github.com/HERA-Team/hera_sim

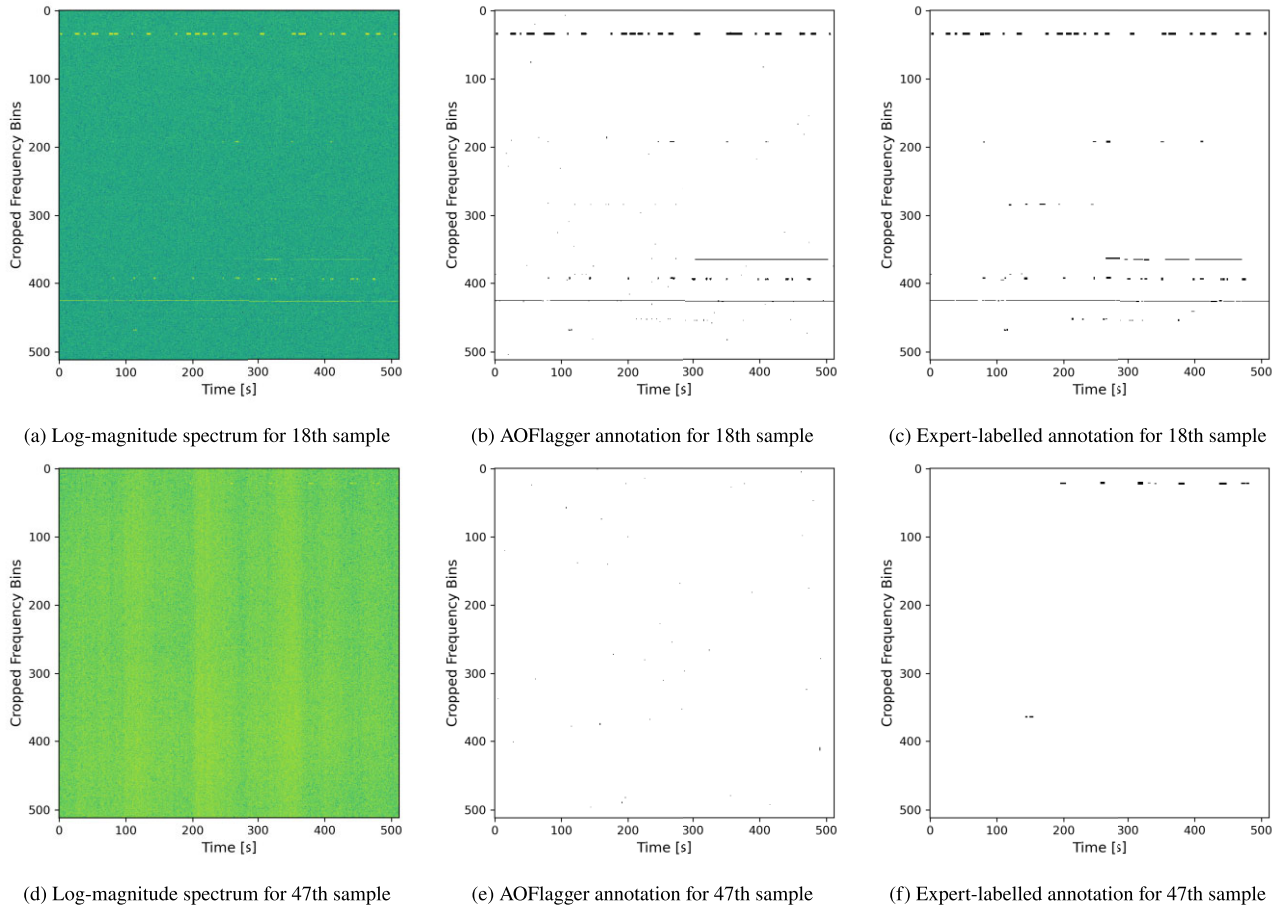


Figure 3. Spectrograms and their masks from the LOFAR data set with the highest overlap between AOFlagger and expert-labelled annotations (top row) and the lowest overlap (bottom row), where similarity is measured using the F1-score.

the precision and recall for a given threshold. For all evaluations across all models in this work, the threshold is fixed to the maximum obtainable F1-score.

In the class-imbalanced scenario of RFI detection, high AUROC scores imply that a model is effective in classifying the majority class. This means that all non-RFI signals are detected as not RFI. Conversely, the AUPRC and F1-scores focus on the minority class, meaning that when AUPRC is high, the model is better at detecting RFI with a low RFI misclassification rate. Therefore, in order to maintain consistency with previous works’ evaluations and to give insight into a model’s performance on the both the majority and minority classes, we evaluate using both AUPRC and AUROC.

We use the AOFlagger masks to train all models in this work. In the case of LOFAR, we use the flags provided by the FLAG field of the measurement sets obtained from the LTA. For the HERA data set, the optimal flagging strategy is determined experimentally. As there is no pre-specified strategy for the HERA telescope, we test the HERA data set on all available strategies for several different base thresholds. We find that the `bihorns-telescope` strategy with a starting threshold of 10 is optimal with respect to the joint maximum of AOFlagger across AUROC, AUPRC, and F1-score as shown in Fig. 4.

For comparison with existing work, we select the state-of-the-art RFI detection models. As described in Section 2, we evaluate with supervised-segmentation algorithms, based on the U-Net (Ronneberger et al. 2015) architecture. These being the magnitude-only

U-Net for RFI detection (Akeret et al. 2017b; Kerrigan et al. 2019) as well as its residual adaptations, R-Net (Sadr et al. 2020) and RFI-Net (Yang et al. 2020). Additionally, we measure the AOFlagger on both data sets and report its performance. We train and evaluate every model three times with a randomly initialized seed and report the mean and standard deviation for each evaluation.

For both data sets, we perform an independent coarse grid search across the hyper-parameters of NLN. These being patch size (n), the number of latent space dimensions (L), number of neighbours (K), maximum latent-distance threshold (T), and discriminative training weight (α). We determine the optimal hyper-parameters based on the average maximum of AUROC and AUPRC. For LOFAR and HERA, we use 16 and 20 neighbours, respectively, and find a patch size of 32×32 to be optimal for both. Furthermore, the latent dimension size of 8 is used for HERA, whereas we find 32 dimensions to be optimal for LOFAR. We threshold the latent-distance vector at its 66th percentile for LOFAR and the 10th percentile for HERA. Finally, we find that the optimal discriminative hyper-parameter α is 0.6 for both the HERA and LOFAR data sets. For all other models, we use all parameters specified by the authors other than the patch size, which we fix to 32×32 in order to keep comparison consistent.

The difference in the hyper-parameters between the data sets is due to the increased complexity of the real LOFAR data relative to the simulated HERA data. This complexity is due to stochastic instrumentation effects, ionospheric artefacts, increased dynamic range, and many more. Therefore, the autoencoder requires a larger

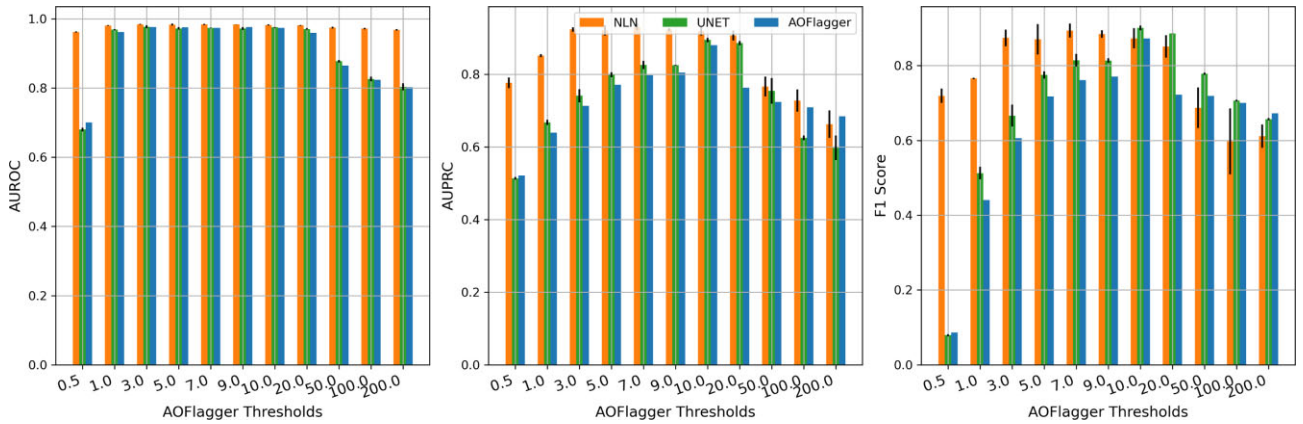


Figure 4. Performance of NLN, U-Net, and AOFlagger on the HERA data. Both U-Net and NLN are trained on the AOFlagger masks for a given starting threshold and evaluated on the simulator ground truth.

Table 2. Performance of RFI detection models on the simulated HERA data set when trained using the AOFlagger annotation at a threshold of 10 and evaluated on the ground truth from the simulator. We do not report standard deviation of AOFlagger as it is deterministic. Best scores are shown in bold.

Model	AUROC	AUPRC	F1-score
AOFlagger (Offringa et al. 2012)	0.9736	0.8799	0.8729
U-Net	0.9746 ± 0.0009	0.8963 ± 0.0074	0.9015 ± 0.0074
RFI-Net (Yang et al. 2020)	0.9728 ± 0.0007	0.8898 ± 0.0086	0.8950 ± 0.0068
R-Net (Sadr et al. 2020)	0.9751 ± 0.0016	0.8456 ± 0.0134	0.8460 ± 0.0223
NLN (ours)	0.9805 ± 0.0017	0.9265 ± 0.0117	0.9103 ± 0.0229

latent dimensionality to better represent this increased complexity. Similarly, an increased latent-distance threshold is used to mitigate the elevated reconstruction error-based noise.

Finally, to validate the suitability of the NLN algorithm for novelty-based RFI detection, we compare it against three commonly used novelty detection techniques. In this case, we consider deep K -nearest neighbours (DKNN; Bergman et al. 2020), autoencoding models with an L2 loss (Bergmann et al. 2019a), and SSIM-based loss (Bergmann et al. 2019b).

5.2 HERA results

In Fig. 4, we illustrate the performance sensitivity of the magnitude-based U-Net and NLN when modifying the AOFlagger starting threshold for the bighorns strategy. It is clear that NLN is less sensitive to changes in the validity of the training data, exhibiting little variation relative to the AOFlagger-based masks. This is because the model is not directly trained on the AOFlagger-based masks. However, U-Net’s performance is shown to be more dependent on the accuracy of the AOFlagger masks, indicating that it indeed is overfitting to the training labels.

For low thresholds (>5), the training data are overflagged, meaning that all RFI is flagged along with a large percentage of the astronomical data. Conversely, for high thresholds (<10), the data are underflagged, meaning that some RFI is not flagged. From this intuition, it is clear that the NLN algorithm is less sensitive to overflagging, but its performance deteriorates when the training data are underflagged. For the experiments using the simulated HERA data set in the remainder of this paper, we fix the AOFlagger threshold to 10, as it gives optimal flagging performance.

In Table 2, we show the performance of NLN for RFI detection relative to the current state of the art. It is clear that the NLN offers

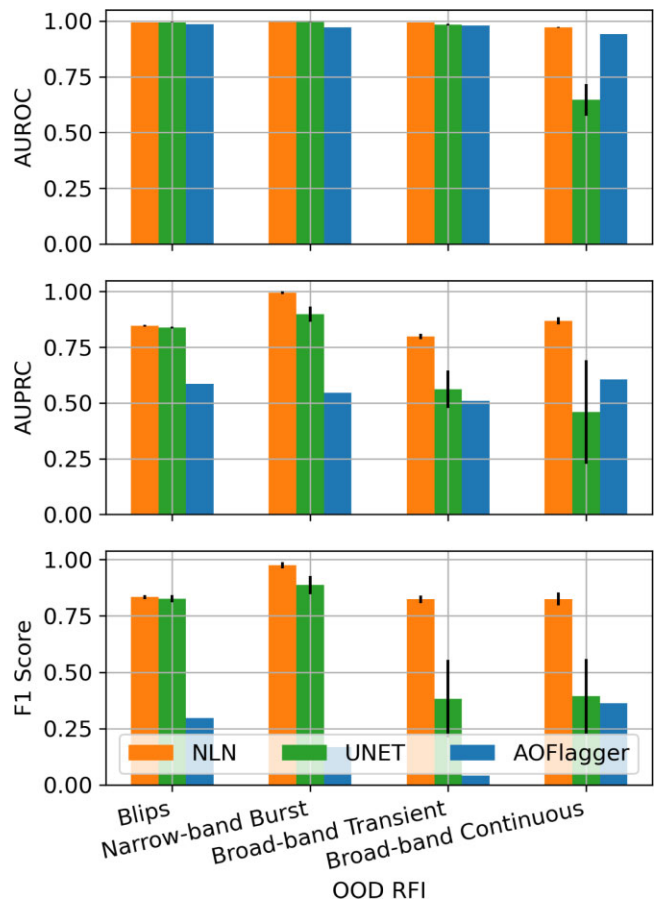


Figure 5. OOD RFI detection performance of the NLN, U-Net, and the AOFlagger. As the RFI is excluded in the training process of NLN, the method is unaffected by OOD effects.

superior performance across all metrics demonstrating the success of our approach to use clean data for training.

In Fig. 5, we demonstrate an interesting unintended consequence of our work in the HERA setting. Here, we train each model on the data set but excluding a particular type of RFI, then during testing we expose the model to the unseen RFI type. This paradigm effectively evaluates how well models generalize to OOD RFI. As the training process of NLN excludes all RFI (OOD or not) from the

Table 3. Performance of RFI detection models on the real LOFAR data set when trained using the AOFlogger annotations from the LTA and evaluated on the expert-labelled ground truth. We do not report AOFlogger and DKNN standard deviations as they are deterministic. Best scores are shown in bold.

Model	AUROC	AUPRC	F1-score
AOFlogger (Offringa et al. 2012)	0.7883	0.5716	0.5698
U-Net	0.8017 ± 0.0058	0.5920 ± 0.0031	0.5876 ± 0.0031
RFI-Net (Yang et al. 2020)	0.8109 ± 0.0037	0.5991 ± 0.0038	0.5979 ± 0.0012
R-Net (Sadr et al. 2020)	0.8301 ± 0.0084	0.5495 ± 0.0145	0.5286 ± 0.0195
AE-L2 (Bergmann et al. 2019a)	0.8397 ± 0.0019	0.3933 ± 0.0036	0.4491 ± 0.0007
AE-SSIM (Bergmann et al. 2019b)	0.7748 ± 0.0046	0.3913 ± 0.0186	0.4801 ± 0.0115
DKNN (Bergman et al. 2020)	0.8285	0.0704	0.1528
NLN (ours)	0.8622 ± 0.0006	0.6216 ± 0.0005	0.5114 ± 0.0004

training set, it does not learn explicit models of the RFI; in effect, our method significantly outperforms both supervised and classical methods across all metrics. This is also clearly demonstrated by the variance across each experiment.

5.3 LOFAR results

In Table 3, we show the performance of NLN relative to the state of the art on the LOFAR data set. Here, we see a similar trend to the HERA-based results; NLN offers superior performance in terms of AUROC and AUPRC. However, in terms of F1-score, RFI-Net offers best performance. The decrease in F1-score is due to the NLN algorithm yielding more false negatives when thresholding all predictions with a single threshold. This is attributable to the large fluctuations in power of the RFI in the LOFAR data set in combination with the reconstruction-error term of the NLN RFI detector. When the RFI power is low, the reconstruction error will be of a low amplitude and NLN will produce predictions with low power. In effect, when these low-power predictions are thresholded using the same threshold as the high-power RFI (in this case one that maximizes F1-score), the outputs have more false negatives. This is in contrast with the HERA data set, which has RFI with a consistently higher power level than the astronomical and system-based signals. A possible solution to this is to predict a threshold on a per spectrogram or patch level; however, we consider this out of scope for this research.

Furthermore, in the bottom half of Table 3, we compare NLN with commonly used novelty detection methods. Here, we see that NLN significantly outperforms these methods on all metrics. We attribute the performance improvement to the combination of reconstruction error and latent error.

Additionally, to test the model’s reliance on data set size, we evaluate each model on a percentage of the training data. We show in Fig. 6 that NLN is less sensitive to reductions in training data size, performing almost uniformly even with large decreases in training data size. Conversely, the supervised methods’ performance scales asymptotically with data set size, exhibiting significantly higher variance in their performance with smaller data set sizes. We associate the asymptotic scaling and increased variance with both the supervised model’s larger capacity (due to residual and skip connections) as well as the diversity of the RFI landscape. An illustration of these effects and a comparison between the models is shown in Fig. 7.

To determine the sensitivity of the parameters of the NLN algorithm, we perform a course grid search of its hyper-parameters. We search across the number of latent dimensions, patch size, and number of neighbours as illustrated in Fig. 8. In order to better visualize the four-dimensional space, we plot cross-sections of the high-dimensional landscape. First, we fix the number of neighbours

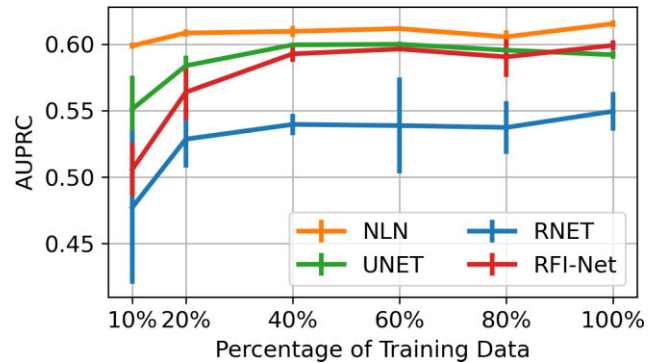


Figure 6. AUPRC performance of each model when training on a percentage of the original LOFAR data set and evaluating on the original LOFAR test data set.

to 16, as shown in Figs 8(a)–(c). It can be seen that the optimal number of latent dimensions is 32, with respect to the average maximum of AUPRC and AUROC. We then set the optimal number of latent dimensions and plot the effect of varying the number of neighbours in Figs 8(d)–(f). Through this, we determine the optimal number of neighbours to be 16 with respect to the average maximum AUROC and AUPRC. We conclude that NLN gives optimal performance when the number of latent dimensions is 32, the patch size is 32×32 , and the number of neighbours is 16.

Finally, as a consequence of our RFI-free selection algorithm, we find that NLN requires 66 per cent less data in comparison to its supervised counterparts. This amount is naturally data set dependent; however, we expect that due to the reduced training data, there will be less compute time and less power consumption while training.

6 CONCLUSIONS

RFI detection is an increasingly important research topic for radio astronomy. State-of-the-art solutions to the RFI problem have been based on supervised machine learning techniques, which fail to address the prohibitive cost of labelling astronomical data. In this work, we have documented how inverting the detection problem effectively addresses this issue. We have shown that NLN provides better than state-of-the-art RFI detection without incurring the cost of labelling.

Furthermore, we have demonstrated that our method better generalizes to unseen RFI, whereas current supervised approaches overfit to weak-label-based RFI masks. As a consequence, we hypothesize that our approach will better generalize to future generations of emitters, whereas existing supervised methods will have to be regularly retrained. Additionally, we find that due to our training patch selection process, we need less data for training, hence decreasing both the training time and energy cost.

This being said, there is sufficient evidence in several other domains that suggests supervised-segmentation algorithms will outperform unsupervised approaches given sufficient high-quality labels. However, in the current state of the RFI detection domain, where there are still a few labelled data sets available and the high cost of obtaining labelled data, we propose the inverted approach as the way forward.

We plan to further improve the performance of NLN applied to RFI detection through additional training priors. Contrastive self-supervised learning is a candidate solution, thanks to its ability to generate more robust latent representations that can be leveraged in the NLN algorithm. Furthermore, in order to improve the increased

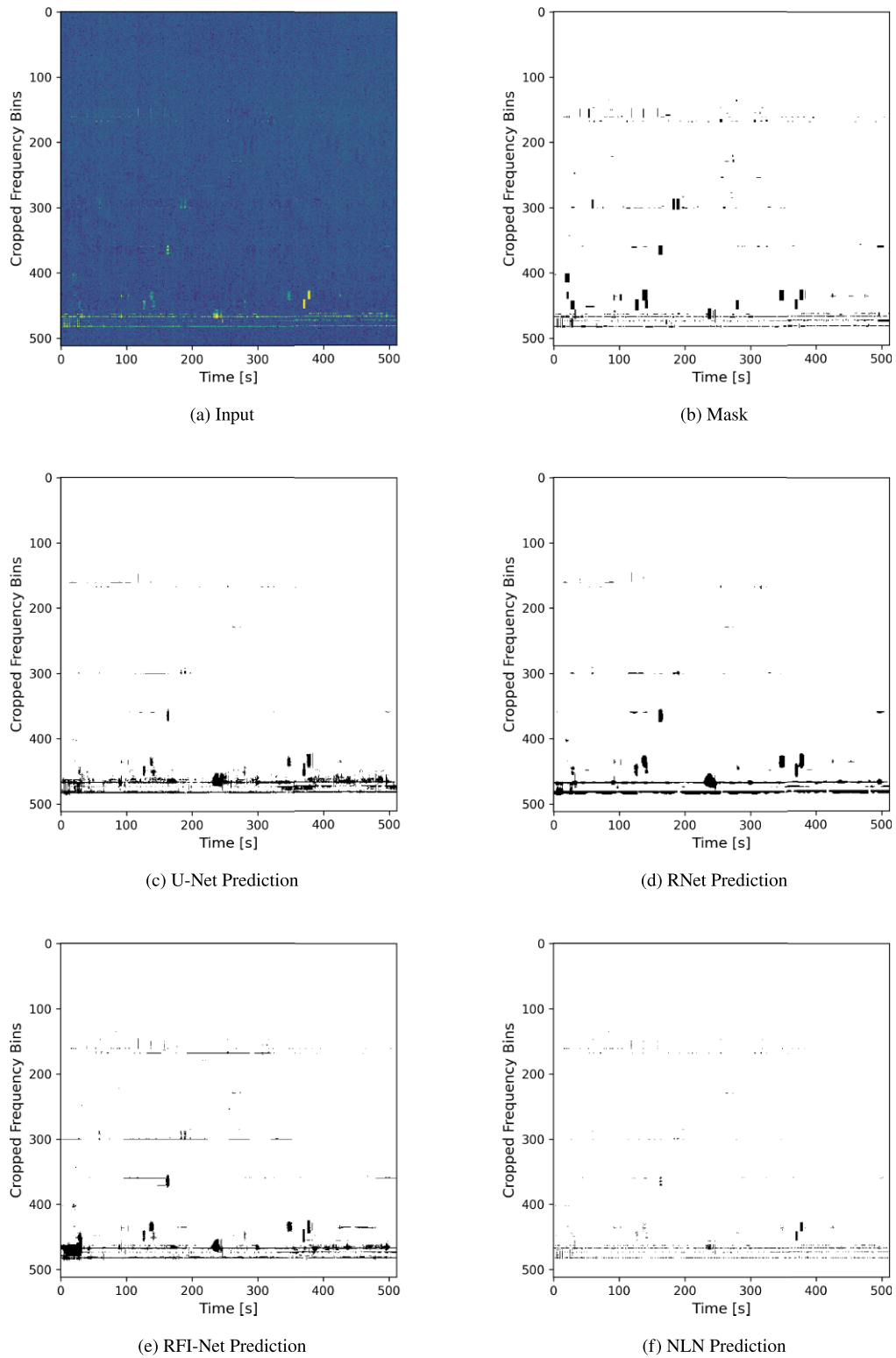


Figure 7. The 95th sample from the testing set, its corresponding mask, and the predictions of each model.

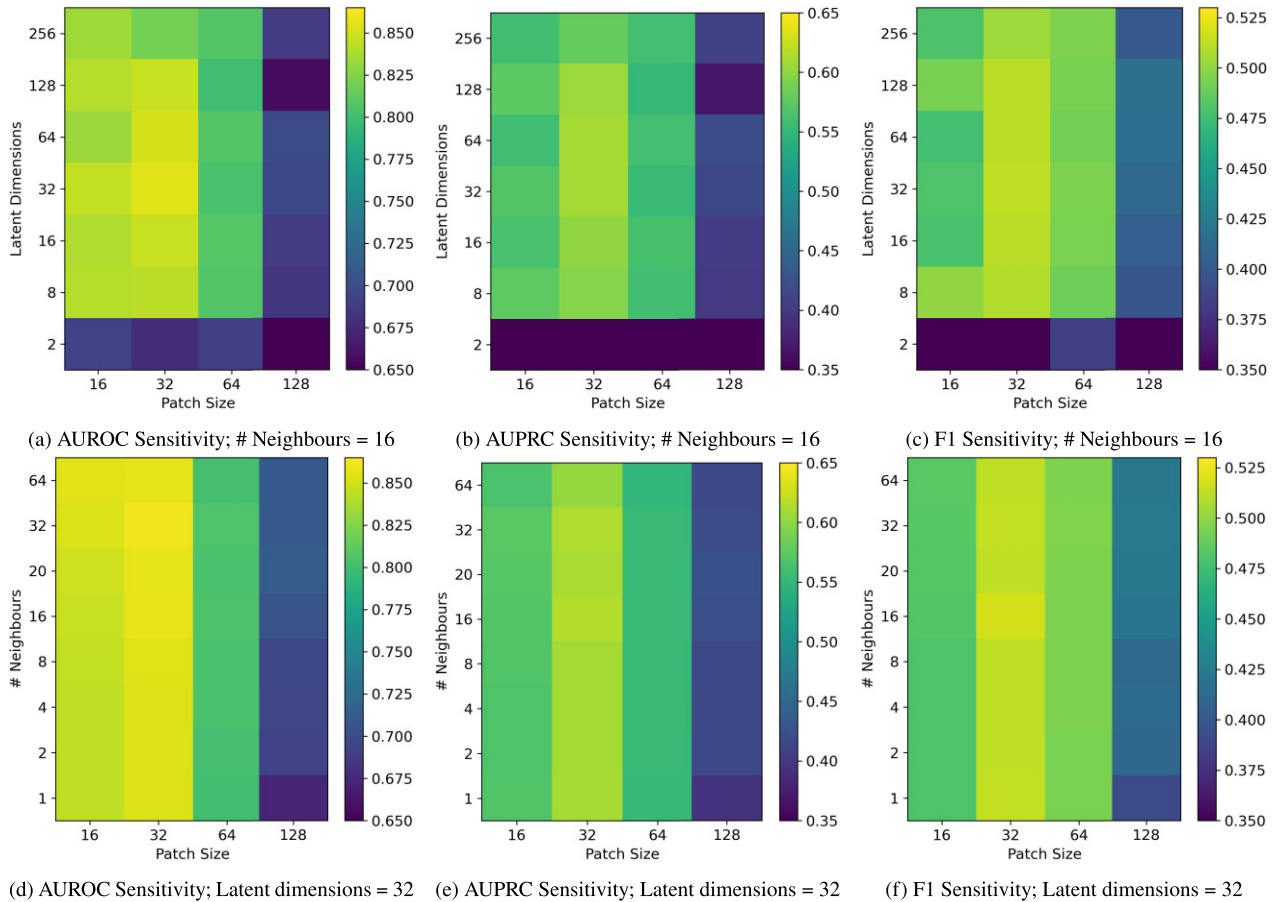


Figure 8. Sensitivity of the hyper-parameters on the LOFAR-based performance of NLN when varying latent dimensionality, patch size, and number of neighbours. In order to visualize the four-dimensional space, the number of neighbours is fixed to 16 in the top row, whereas in the bottom row, we fix the number of latent dimensions to 32. The optimal parameters for the LOFAR data set are a patch size of 32×32 , 32 latent dimensions, and 16 neighbours.

false negative rate of NLN on the LOFAR data set, we suggest more research to be done into automated processing schemes to deal with the high dynamic range of astronomical data when training unsupervised models. This could be additionally improved using a hybrid approach through `SumThreshold` or trying to directly predict a threshold on a per patch basis. Finally, we plan to extend this work to more general anomaly detection-based problems within radio astronomy.

ACKNOWLEDGEMENTS

This work is part of the ‘Perspectief’ research programme ‘Efficient Deep Learning’ (EDL, <https://efficientdeeplearning.nl>), which is financed by the Dutch Research Council (NWO) domain Applied and Engineering Sciences (TTW). The research makes use of radio astronomy data from the LOFAR telescope, which is operated by ASTRON (Netherlands Institute for Radio Astronomy), an institute belonging to the Netherlands Foundation for Scientific Research (NWO-I). Special thanks to Prof. dr. ir. Cees T. A. M. de Laat for useful advice and feedback on this research.

DATA AVAILABILITY

The data underlying this paper are available on Zenodo, at <https://dx.doi.org/10.5281/zenodo.6724065>

REFERENCES

- Akçay S., Atapour-Abarghouei A., Breckon T. P., 2019, in *Computer Vision – ACCV 2018*. Springer International Publishing, p. 622
- Akeret J., Seehars S., Chang C., Monstein C., Amara A., Refregier A., 2017a, *Astron. Comput.*, 18, 8
- Akeret J., Chang C., Lucchi A., Refregier A., 2017b, *Astron. Comput.*, 18, 35
- Baan W. A., Fridman P. A., Millenaar R. P., 2004, *ApJ*, 128, 933
- Bergman L., Cohen N., Hoshen Y., 2020, preprint ([arXiv:2002.0445](https://arxiv.org/abs/2002.0445))
- Bergmann P., Fauser M., Sattlegger D., Steger C., 2019a, in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*. IEEE, p. 9584
- Bergmann P., Löwe S., Fauser M., Sattlegger D., Steger C., 2019b, *VISIGRAPP 2019 – Proc. 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, Vol. 5, p. 372
- Breunig M. M., Kriegel H.-P., Ng R. T., Sander J., 2000, *ACM SIGMOD Record*, 29, 93
- Chang C., Monstein C., Akeret J., Seehars S., Refregier A., Amara A., Glauser A., Stuber B., 2017, *MNRAS*, 464, 1727
- DeBoer D. R. et al., 2017, *PASP*, 129, 45001
- Farahani A., Voghohi S., Rasheed K., Arabnia H. R., 2021, *Advances in Data Science and Information Engineering*, 877
- Foley A. R. et al., 2016, *MNRAS*, 460, 1664
- Fridman P. A., Baan W. A., 2001, *A&A*, 378, 327
- Hainaut O. R., Williams A. P., 2020, *A&A*, 636, 1
- Harrison S., Coles R., Robshaw T., Del Rizzo D., 2019, in *2019 RFI Workshop – Coexisting with Radio Frequency Interference (RFI)*, IEEE, p. 1

- Johnson J., Douze M., Jegou H., 2021, *IEEE Trans. Big Data*, 7, 535
- Kerrigan J. et al., 2019, *MNRAS*, 11, 1
- Kildal P.-S., Baker L., Hagfors T., 1991, *IEEE Antennas Propag. Mag.*, 33, 12
- Kingma D. P., Ba J. L., 2015, 3rd International Conference on Learning Representations, ICLR 2015 – Conference Track Proceedings, p. 1
- Larsen A. B. L., Sønderby S. K., Larochelle H., Winther O., 2016, Proceedings of the 33rd International Conference on International Conference on Machine Learning - Volume 48, JMLR.org, New York, NY, USA, p. 1558
- Lochner M., Bassett B., 2021, *Astron. Comput.*, 36, 100481
- Malanchev K. L. et al., 2021, *MNRAS*, 502, 5147
- Margalef-Bentabol B., Huertas-Company M., Charnock T., Margalef-Bentabol C., Bernardi M., Dubois Y., Storey-Fisher K., Zanisi L., 2020, *MNRAS*, 496, 2346
- Maslakovic S., Linscott I., Osllick M., Twicken J., 1998, in Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '98 (Cat. No.98CH36181) vol. 3, IEEE, Seattle, WA, USA, p. 1793
- Mesarcik M., Boonstra A.-J., Meijer C., Jansen W., Rangelova E., van Nieuwpoort R. V., 2020, *MNRAS*, 496, 1517
- Mesarcik M., Rangelova E., Boonstra A.-J., van Nieuwpoort R. V., 2022, *Array*, 14, 100182
- Offringa A. R., de Bruyn A. G., Biehl M., Zaroubi S., Bernardi G., Pandey V. N., 2010, *MNRAS*, 405, 155
- Offringa A. R., Van De Gronde J. J., Roerdink J. B., 2012, *A&A*, 539, A95
- Ronneberger O., Fischer P., Brox T., 2015, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9351, p. 234
- Sadr A. V., Bassett B. A., Oozeer N., Fantaye Y., Finlay C., 2020, *MNRAS*, 499, 379
- Schlegl T., Seeböck P., Waldstein S. M., Schmidt-Erfurth U., Langs G., 2017, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10265, p. 146
- Sokolowski M. et al., 2015, *PASA*, 32, e004
- Srivastava A., Valkov L., Russell C., Gutmann M. U., Sutton C., 2017, *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Curran Associates Inc., Long Beach, California, USA, p. 3310
- Tang H., Scaife A. M., Leahy J. P., 2019, *MNRAS*, 488, 3358
- Tony Liu F., Ming Ting K., Zhou Z.-H., 2008, *2008 Eighth IEEE International Conference on Data Mining*, Pisa, Italy, p. 413
- van Haarlem M. P. et al., 2013, *A&A*, 556, A2
- Villar V. A., Cranmer M., Berger E., Contardo G., Ho S., Hosseinzadeh G., Lin J. Y.-Y., 2021, *ApJS*, 255, 24
- Vinsen K., Foster S., Dodson R., 2019, *2019 URSI Asia-Pacific Radio Science Conference (AP-RASC)*, New Delhi, India, p. 2
- Vos E. E., Francois Luus P. S., Finlay C. J., Bassett B. A., 2019, *2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP)*, Pittsburgh, PA, USA, p. 1
- Yang Z., Yu C., Xiao J., Zhang B., 2020, *MNRAS*, 492, 1421
- Zhang Y. G., Hyun Won K., Son S. W., Siemion A., Croft S., 2018, *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP 2018)*, Anaheim, CA, USA, p. 1114

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.