

Radio frequency interference detection based on the AC-UNet model

Rui-Qing Yan¹, Cong Dai¹, Wei Liu², Ji-Xia Li³, Si-Ying Chen¹, Xian-Chuan Yu^{1*}, Shi-Fan Zuo³ and Xue-Lei Chen³

¹ College of Information Science and Technology, Beijing Normal University, Beijing 100875, China; yuxianchuan@163.com, 201822210015@mail.bnu.edu.cn

² National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences; Beijing 100190, China; liuwei@ia.ac.cn

³ National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China; xuelei@nao.cas.cn, jxli@bao.ac.cn

Received 2020 September 30; accepted 2020 December 3

Abstract Radio frequency interference (RFI) is a serious issue in radio astronomy. This paper proposes a U-Net network model with atrous convolution to detect RFI. Using the ability of convolutional neural networks to extract image features of RFI, and learning RFI distribution patterns, the detection model of the RFI is established. We use observational data containing real RFIs obtained by the Tianlai telescope to train the model so that the model can detect RFI. Calculate the probability of a data point being RFI pixel by pixel, and set a threshold. At the same time the dropout layer was added to avoid overfitting problems. If the predicted probability of a data point exceeds the threshold, it is considered that there is RFI, and if the predicted probability of a data point does not exceed the threshold, then it is considered that there is no RFI, so that the part of the image with RFI is flagged. Experimental results show that this approach can achieve satisfactory accuracy in the detection of radio observation images with a small amount of RFI.

Key words: methods: data analysis — techniques: image processing — methods: observational

1 INTRODUCTION

In radio astronomy, radio frequency interference (RFI) broadly refers to the influence of human communication activities and natural interference signals that affect the reception of weak astronomical signals by radio telescopes (Akeret et al. 2017; Lahtinen et al. 2017). RFI restricts the search for these weak pulse signals, and even affects the task of searching for celestial bodies such as radio transient sources. With the continuous development of astronomical research and technology, the sensitivity of radio astronomical equipment has increased, allowing astronomical researchers to make observations on a wider frequency range. But at the same time, with the rapid development of human communication technology, signals and noises generated by various human activities occupy more and more frequency bands, and the influence range is getting wider and wider, which has a serious impact on astronomical observation data. Therefore, it is very important to effectively process the complex radio astronomical observation data and perform RFI detection.

In the time domain, RFI can be stable, oscillating, repetitive and/or isolated, and may have a wide time scale. In the frequency domain, RFI may be a curved band, wide band, narrow band, regular structure and/or irregular structure. The mixing of time and frequency modes is also common, such as frequency sweep signals (An et al. 2017). Figure 1 shows the wideband and narrowband RFI shown in the LOFAR observation data (Offringa et al. 2013). RFI may be instantaneous or long-lasting. For instantaneous RFI, corresponding methods should be used to separate it from the signals generated by actual transient celestial sources (such as radio transient sources) (An et al. 2017). In addition to this, RFI may occur in the form of pulsed bursts (high amplitude and short time interval), lasting a long time or very short time (Zhu et al. 2017). In order to ensure the quality of astronomical observation data, there are many methods for detecting and eliminating RFI. Filtering technology is common in the field of interference cancellation. Kocz et al. (2010) applied spatial filtering to the actually acquired multi-beam data. The spatial filter can use the relative arrival time of signals on multiple sensors to identify and separate signals

* Corresponding author

from different directions. These samples are classified by the relevant characteristics, and then the RFI is eliminated, which proves the effectiveness of the filtering technology. The noise elimination methods based on nonlinear filtering mainly include singular value decomposition (SVD) method. SVD can map the noisy signal space to the no-noise signal space, and perform singular value decomposition on the signal to achieve the purpose of suppressing noise and eliminating RFI (Fridman & Baan 2001). Methods based on linear transformation mainly include principal component analysis (PCA), independent component analysis (ICA), factor analysis (FA), etc. Czech et al. (2017) performed statistical characterization of RFI signals and studied the use of PCA and nuclear PCA and other component analysis methods to detect RFI. In addition, Offringa et al. (2010) uses the “SumThreshold” method, which treats the time-frequency plane as a two-dimensional image, and eliminate the RFI using image processing techniques. The filtering technique depends on the selection of the filter. The method is simple, but the effect is poor, and the methods such as PCA cannot deal well with the nonlinear problem, and for massive observation data, strong data feature extraction is a required ability to deal with the model.

The rapid development of deep learning (Deng & Yu 2014; Schmidhuber 2015; Sainath et al. 2015) technology has brought great breakthroughs in many fields, such as medicine (Greenspan et al. 2016), education (Greenspan et al. 2016), finance (Heaton et al. 2016) and driverless driving (Fayjie et al. 2018). Computer vision is an important direction in the field of artificial intelligence. It includes computer vision tasks such as image segmentation (He et al. 2017; Badrinarayanan et al. 2018; Chen et al. 2017b), image classification (Xia et al. 2017; He et al. 2016; Simonyan & Zisserman 2014) and object detection (Ren et al. 2015; Redmon et al. 2016; Liu et al. 2016). Because deep learning can extract the deep features of images, deep learning technology has a good effect in the realization of these visual tasks. Although image semantic segmentation is a hot topic in the field of computer vision, there are few studies on semantic segmentation combining the characteristics of data in astronomy and specific scientific needs. The image semantic segmentation method can classify the image at the pixel level, that is, classify each pixel in the image, and then obtain the boundary between different subjects in the image, complete the segmentation process, and also get the mark of the subject. In the task of RFI detection, RFI is often flagged in the form of points and lines. We also hope to get the location of RFI and get the corresponding flag. The traditional flagging method or flagging based on human experience is inefficient. In addition, automatic and efficient data processing methods are needed in massive

data astronomical observations, especially for large area sky surveys or high time resolution observations.

Based on the above analysis, in order to accurately and comprehensively perform RFI detection on images with a small amount of RFI, improve the efficiency and recall rate of RFI detection, and be able to effectively learn the characteristics of RFI images, this paper proposes an AC-UNet model with atrous convolution to detect RFI in the real astronomical observation data, and calculate the probability of a pixel being RFI in the image pixel by pixel. At the same time, the dropout layer is added to avoid overfitting problems, and we set a threshold to flag the part of the image that has RFI.

Our paper is structured as follows. The principle of atrous convolution is introduced in Section 2. In Section 3, we describe the data and preprocessing method. In Section 4, we introduce the structure of the RFI detection model. Section 5 discusses experimental results obtained from the model proposed in Section 4. Finally, conclusions are presented in Section 6.

2 ATRous CONVOLUTION PRINCIPLE

Atrous convolution (Chen et al. 2014), also known as dilation convolution, is born in the field of image segmentation and can control the calculation of feature density in fully convolutional networks. By setting the expansion coefficient of the convolution kernel, filling the convolution kernel with 0 (Atrous) to expand the size of the convolution kernel, and then performing the convolution operation on the image, the effect of increasing the receptive field without changing the size of the output image is achieved. The receptive field is a part of the original input image that should correspond to a node in the output feature map. Expanding the receptive field can learn the original image features more comprehensively and carefully. DeepLabV2 uses the method of Atrous Spatial Pyramid Pooling (ASPP) (Chen et al. 2017a) to expand the original convolution kernel into multiple-size convolution kernels with atrous convolution to obtain features of different sizes, and then fuses the features to get the final prediction result. Figure 2 is an example of the expansion of convolution kernels corresponding to different expansion coefficients in ASPP. It can be seen that the larger the coefficient is, the larger the gap between two adjacent values in the convolution kernel is.

There are two implementation forms of atrous convolution in the keras framework to expand the receptive field. In the first form, the atrous coefficient `dilation_rate` is directly set in the ordinary convolution `Convolution2D` layer. In the second form, the `AtrousConvolution2D` layer is used in the keras framework, and the convolution kernel expansion coefficient is set in the parameters. Take

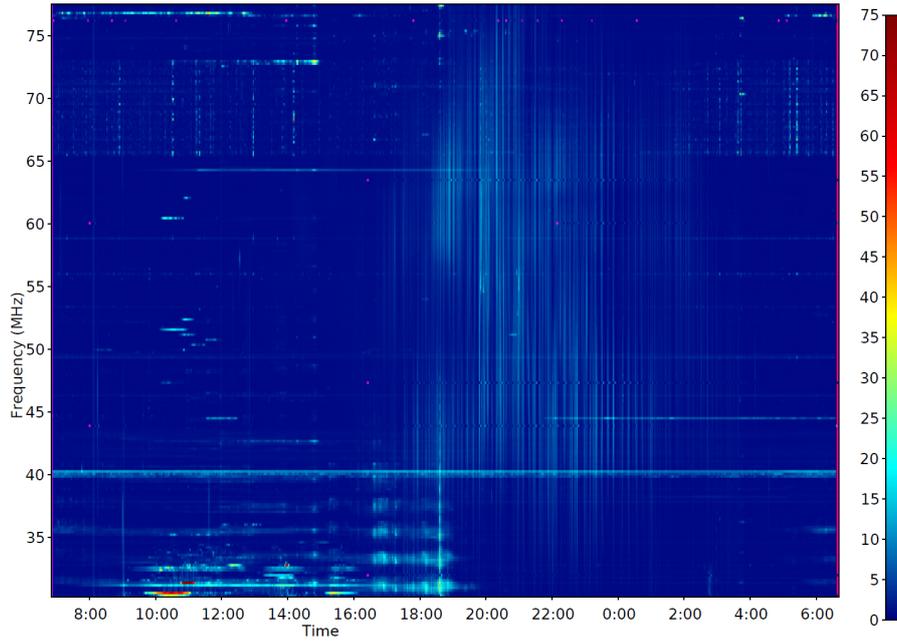


Fig. 1 Dynamic RFI occupancy spectrum for low-band antennas (LBA) surveys. Color intensity represents RFI occupancy, from 0% to 100% (Offringa et al. 2013).

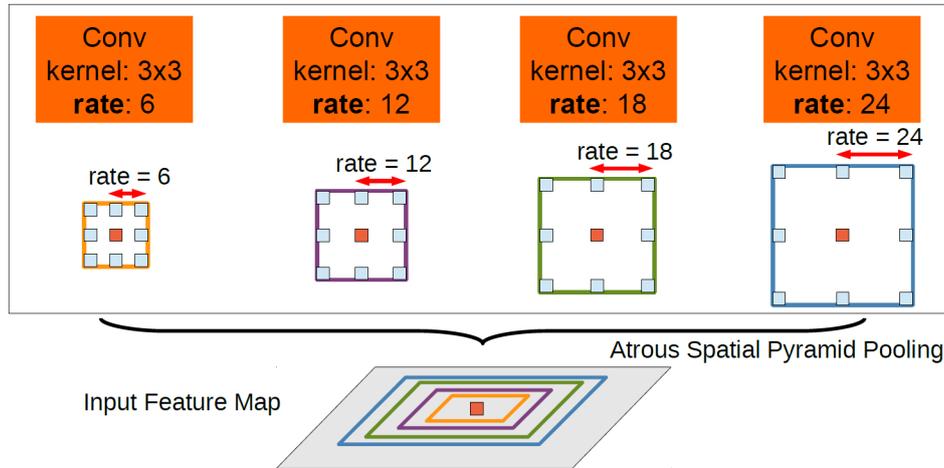


Fig. 2 ASPP schematic. To classify the center pixel (orange), ASPP exploits multi-scale features by employing multiple parallel filters with different rates. The effective field-of-views are shown in different colors (Chen et al. 2017a).

an example, for a convolution kernel filter of the size 3×3 , the convolution kernel expansion coefficient is set to be $\text{atrous_rate}=2$, and $(\text{atrous_rate}-1)$ zeros are inserted between two adjacent values in the original convolution kernel. In this way, the size of the convolution kernel filter is expanded from 3×3 to 5×5 after $3 + (3 - 1) \times (2 - 1) = 5$ operations. The expanded convolution kernel is shown in Figure 3. The size of the convolution kernel, filter_heights^* , is calculated in Equation (1) (Chen et al. 2017a).

$$\text{filter_height}^* = \text{filter} + (\text{filter} - 1) \times (\text{atrous_rate} - 1). \quad (1)$$

3 DATA AND PREPROCESSING

3.1 Data

In this work, we analyzed the real observation data from the Tianlai experiment (Chen et al. 2015), a radio telescope array dedicated to the neutral hydrogen sky survey over a frequency range of 400 MHz to 1420 MHz. The array is located in the Hongliuxia station, a very radio-quiet area in Xinjiang, China (Wu et al. 2014). Currently, Tianlai, in its Pathfinder stage, has built two types of antennas – cylinder and dish – both working in the frequency range 700–800 MHz. The data in this work is obtained by the

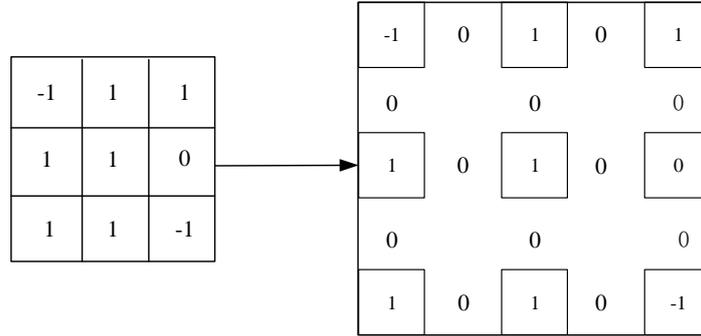


Fig. 3 Convolution kernel expansion process when the expansion coefficient is 2. Left: the original convolution kernel. Right: the expanded convolution kernel.

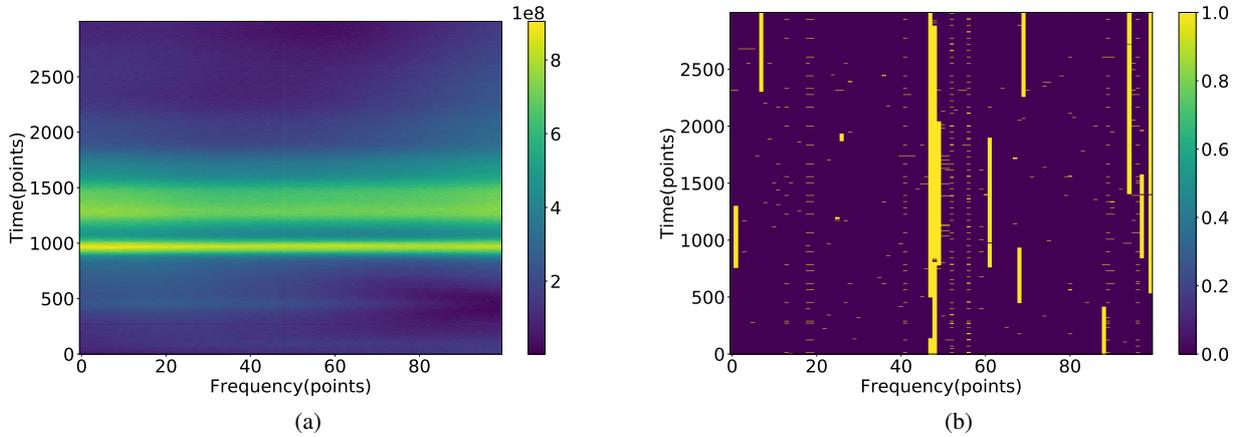


Fig. 4 Observation data of a typical baseline taken by Tianlai cylinder array. (a) The amplitude of the observation data. (b) The corresponding RFI flag.

cylinder array. Preliminary analysis shows that the site has suffered from a few single frequency RFIs generated by the digital devices in the first year's running. Later on, the devices were upgraded to avoid RFI leakage, and data analysis work using observation data in subsequent years sees almost no RFI (Li et al. 2020). To better train the model and check the RFI detection performance, we will use the data taken in the first year as a trial.

The observation data are taken between 20:15:45 Sep. 27th and 00:18:45 Sep. 28th Beijing time (UTC+08h) in 2016. The integration time is 4 seconds, so there are 3645 time points. The frequency dimension has 100 frequency points, corresponding to a frequency range of 744 MHz to 756 MHz. Tianlai cylinder has 96 dual linear polarization feeds, each producing two polarization signals. The correlation of any two signals is the so-called visibility. For Tianlai cylinder, the whole correlation pairs are therefore 18528. The visibility is actually a complex value: its amplitude reflects the received power level by the antenna and its phase is related to the direction of the incoming radio signal. In the analysis of RFI detection, we only focus on the amplitude part which

is obtained by calculating the modulus of the complex value. Figure 4(a) shows the observation data of a typical baseline, the larger the amplitude is, the brighter the color is displayed in the image. Figure 4(b) is the RFI flag corresponding to the observation data of a typical baseline.

3.2 Preprocessing

In the actual observation of the Tianlai telescope, the instrumental effects are traced by a calibrator noise source (CNS) which regularly emits a wideband signal. Therefore, before inputting the network model, these CNS data are firstly removed, and the total number of time points which contains celestial signals is reduced to 3340. Each correlation pair corresponds to a time-frequency two-dimensional plane (3340×100), which we use as a picture and input to our network model. In addition, since the number of frequency points of the observation data in this batch is 100, in order to facilitate calculation and observe the data points at each time and frequency in more detail, we divide a whole picture into several (100×100) data blocks. So in the end, 200 baseline time-frequency data

and corresponding RFI flag data were selected as training data and test data. Among them, 4800 data blocks with a size of (100×100) and RFI flags corresponding to each data block are selected as training data to train the model, and 1200 data blocks with a size of (100×100) are used as test data.

After preparing the training data and test data, considering the large amplitude fluctuation range, in order to facilitate data processing and faster convergence when the model is running, this experiment uses the Min-Max standardized method to normalize the data and map the amplitude to the interval $[0,1]$. The normalization formula of minimum and maximum is defined as follows:

$$x^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}}. \quad (2)$$

Here, x_{\max} is the maximum value in the sample data, and x_{\min} is the minimum. Through this calculation formula, each x^* value is in the interval $[0,1]$.

4 MODEL ARCHITECTURE

4.1 AC-Unet Model

This section mainly introduces the AC-Unet model for RFI detection based on the U-Net (Ronneberger et al. 2015) network structure and a detailed description of the model structure. U-Net is a deep convolutional neural network with a U-shaped structure based on a fully convolutional neural network, which can better restore the details of the image. Including the contraction path and expansion path, the image features are extracted through the convolution layer, including the position information. The experiment in this paper designs and builds a 21-layer AC-Unet model based on the size of the input data (100×100) , as shown in Figure 5. The model includes the following layers:

The model includes 14-layer atrous convolution layers (AC layer). In Section 2, the role of atrous convolution is introduced. In order to expand the receptive field, more detailed segmentation results are obtained. In the experiment, two methods of atrous convolution were realized, which verified the effectiveness of atrous convolution. Except the activation function used by the last convolutional layer is the Sigmoid function, the activation functions used by other convolutional layers are all ReLU functions.

The model includes two-layer maximum pooling layers. The pooling layer in the contraction path uses the maximum pooling method to extract the salient features in an area. The pooling size is 2×2 and the step size is 2, which is equivalent to 2 times downsampling, that is, 100×100 data block dimensionality reduction twice to 25×25 . The model also includes two-layer fusion layers. Combining the features extracted by the convolutional

layer of the contraction path and the output results of the upsampling layer corresponding to the expansion path, and fusing the information lost in the feature extraction into the expansion path, can make the output result more refined. For this experimental data, U-Net can merge two pooling results (Pool1, Pool2).

The model includes one-layer dropout layers. Make the model have better prediction results on the test set, and use the dropout layer in AC-Unet model to accelerate the model convergence. The model also includes two-layer upsampling layers. In order to obtain the output of the same size as the original input, in the expansion path, the features of the same size extracted by the corresponding contraction layer are merged, up-sampling is performed twice, and the 25×25 data block is up-sampled to the original size of 100×100 .

4.2 Loss Function and Algorithm

The loss function used in the experiment is the cross-entropy loss function, the formula is as follows:

$$\begin{aligned} L(Y, P(Y/X)) &= -\log P(Y/X) \\ &= -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^M y_{ij} \log(p_{ij}). \end{aligned} \quad (3)$$

Here, L is the loss function, Y is the output variable, and X is the input variable. N is the input sample size, M is the number of possible categories, and y_{ij} is a binary index indicating whether category j is the real category of input instance x_i . p_{ij} is the probability that the model prediction input instance x_i belongs to category j . By minimizing the loss function, the best model weights and bias parameters are obtained, so that the model can more accurately predict the probability value of RFI at each data point.

A suitable optimization algorithm can accelerate model convergence, learn data features more accurately, better adjust the weight and bias parameters of the neural network, and minimize the loss function to the greatest extent. Therefore, this paper considers that the amount of RFI in the experimental data accounts for a relatively low proportion of the total number of data points, so it is relatively sparse. For relatively sparse data sets, the method of using adaptive learning rate will be more conducive to the accuracy of model training and prediction; moreover, the model network is deeper and hopes to converge faster. Therefore, in this experiment, the Adaptive Moment Estimation (Adam) (Kingma & Ba 2014) algorithm is selected to minimize the loss function of the AC-Unet model. Details of the Adam algorithm are given in Table 1.

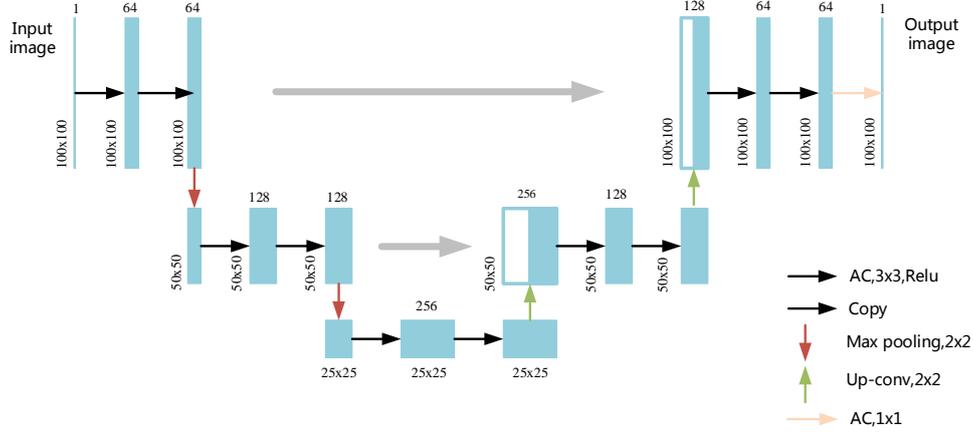


Fig. 5 AC-UNet network structure.

Algorithm 1 Adam, a stochastic optimization algorithm; when the parameters do not converge, we use the Adam algorithm to minimize the loss function of the AC-UNet model.

Require: α Stepsize;

Require: $\beta_1, \beta_2 \in [0, 1)$ Exponential decay rates for the moment estimates;

Require: $f(\theta)$: Stochastic objective function with parameters θ ;

Require: θ_0 : Initial parameter vector;

$m_0 \leftarrow 0$ (Initialize 1st moment vector);

$v_0 \leftarrow 0$ (Initialize 2nd moment vector);

$t \leftarrow 0$ (Initialize timestep);

1: **while** θ_t not converged **do**

2: $t \leftarrow t + 1$

3: $g_t \leftarrow \nabla_{\theta} f_t(\theta_{t-1})$ (Get gradients w.r.t. stochastic objective at timestep t)

4: $m_t \leftarrow \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t$ (Update biased first moment estimate)

5: $v_t \leftarrow \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2$ (Update biased second raw moment estimate)

6: $\hat{m}_t \leftarrow m_t / (1 - \beta_1^t)$ (Compute bias-corrected first moment estimate)

7: $\hat{v}_t \leftarrow v_t / (1 - \beta_2^t)$ (Compute bias-corrected second raw moment estimate)

8: $\theta_t \leftarrow \theta_{t-1} - \alpha \cdot \hat{m}_t / (\sqrt{\hat{v}_t} + \epsilon)$ (Update parameters)

9: **return** θ_t (Resulting parameters).

5 EXPERIMENTAL ANALYSIS

This section mainly introduces the experimental results of the AC-Unet model proposed in this paper. Figure 6 is the given RFI flag. The yellow part indicates the location of RFI, the corresponding flag value is “True”, and the position flag value without RFI is “False”. We use the recall rate and F1-score to measure the classification effect on the existence and nonexistence of RFI respectively; use the AUC value to evaluate the comprehensive classification effect of the model, and use the confusion matrix and ROC curve to visually display the classification effect of each model.

In order to obtain the most suitable U-Net model on the data set, but also to improve the efficiency of RFI detection and various measurement indicators, this section sets up different comparative experiments. We practiced the two implementations of atrous convolution introduced in Section 2, and set different coefficients at the same time to find a more appropriate degree of expansion of the convolution kernel. We also made

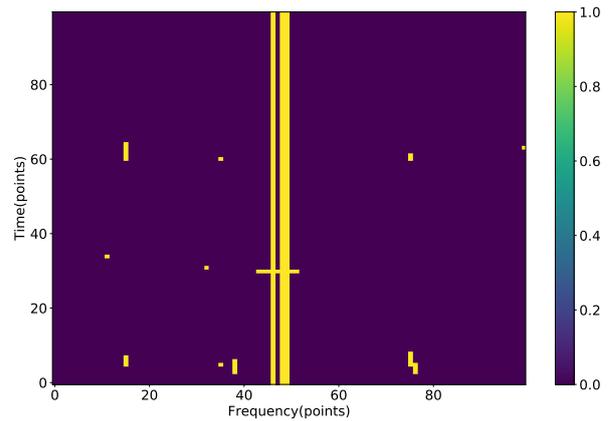


Fig. 6 Data block 694 raw RFI flag.

an experimental comparison with the Full Convolutional Network (FCN) (Long et al. 2015). The experimental results prove that for images with fewer RFI data points, the AC-UNet model can recall more data points with RFI. The different methods used in the experiment are

Table 1 The Amount of RFI in the Test Sample

No.	False Number	True Number
694	9668	332
518	9599	401
510	9521	479
512	9515	485
1087	9250	750

Table 2 Experimental Results of Various Models on Data Block 694

Network Model	Classification	Recall Rate	F1-Score	AUC
U-Net	False	0.98	0.99	0.89
	True	0.81	0.66	
	avg/total	0.97	0.97	
DC-UNet dilation_rate=3	False	0.99	0.99	0.77
	True	0.56	0.63	
	avg/total	0.98	0.98	
AC-UNet atrous_rate=3	False	0.97	0.99	0.94
	True	0.90	0.68	
	avg/total	0.97	0.98	
DC-UNet dilation_rate=4	False	0.98	0.99	0.94
	True	0.90	0.72	
	avg/total	0.98	0.98	
AC-UNet atrous_rate=7	False	0.98	0.99	0.94
	True	0.91	0.75	
	avg/total	0.98	0.98	
AC-FCN	False	0.98	0.99	0.87
	True	0.76	0.66	
	avg/total	0.97	0.98	

described below: (1) DC-UNet dilation_rate=3: Set the atrous coefficient to 3 in the Convolution2D layer in the original U-Net model. (2) AC-UNet atrous_rate=3: Use the atrous convolutional layer AtrousConvolution2D and an expansion coefficient of 3. (3) DC-UNet dilation_rate=4: Set the atrous coefficient to 4 in the Convolution2D layer in the original U-Net model. (4) AC-UNet atrous_rate=7: Use the atrous convolutional layer AtrousConvolution2D and an expansion coefficient of 7.

We use the same training data to train each model that is set in the experiment, and predict the same test data after obtaining the training model. We use data blocks 694, 510, 512 with less than 500 RFI, and randomly select the results of two other test samples 518, 1087 data blocks for display. Similarly, in order to achieve point-by-point prediction, the last layer of the model is a $1 * 1$ convolution layer, and the activation function used is the Sigmoid function. So the output value is the probability value of the existence of RFI predicted for each data point. We set the threshold to 0.5, and mark the value greater than 0.5 as "1", otherwise mark as "0". The amount of RFI contained in the test sample is shown in Table 1; Table 2 to Table 6 show the measurement index values of the various network models on the five test samples.

For the data block 694, only 332 of the 10000 data points have RFI, and they are concentrated around the

Table 3 Experimental Results of Various Models on Data Block 518

Network Model	Classification	Recall Rate	F1-Score	AUC
U-Net	False	0.99	0.98	0.75
	True	0.52	0.57	
	avg/total	0.97	0.97	
DC-UNet dilation_rate=3	False	0.98	0.98	0.80
	True	0.62	0.60	
	avg/total	0.97	0.97	
AC-UNet atrous_rate=3	False	0.97	0.98	0.84
	True	0.71	0.59	
	avg/total	0.96	0.96	
DC-UNet dilation_rate=4	False	0.97	0.98	0.86
	True	0.77	0.60	
	avg/total	0.96	0.96	
AC-UNet atrous_rate=7	False	0.96	0.98	0.89
	True	0.81	0.60	
	avg/total	0.96	0.96	
AC-FCN	False	0.97	0.97	0.74
	True	0.50	0.45	
	avg/total	0.95	0.95	

Table 4 Experimental Results of Various Models on Data Block 510

Network Model	Classification	Recall Rate	F1-Score	AUC
U-Net	False	0.98	0.99	0.92
	True	0.88	0.75	
	avg/total	0.97	0.97	
DC-UNet dilation_rate=3	False	0.98	0.98	0.93
	True	0.89	0.75	
	avg/total	0.97	0.97	
AC-UNet atrous_rate=3	False	0.97	0.98	0.91
	True	0.86	0.68	
	avg/total	0.96	0.96	
DC-UNet dilation_rate=4	False	0.96	0.99	0.92
	True	0.89	0.66	
	avg/total	0.96	0.96	
AC-UNet atrous_rate=7	False	0.99	0.99	0.94
	True	0.90	0.87	
	avg/total	0.99	0.99	
AC-FCN	False	0.97	0.98	0.93
	True	0.89	0.73	
	avg/total	0.97	0.97	

band of 40-50MHz. Figures 7 to 9 show the prediction results of each model on the 694 data block (Fig. 6) for the location of RFI. Figure 7(a) shows the prediction results by our proposed model on 694 data block. Comparing the 694 data block flags shown in Figure 6, we can find that our model predicts the main RFIs in the band of 40–50 frequency points.

First consider the case of setting the dilation_rate value in the convolutional layer. It can be seen from the above figure that the prediction index and image restoration degree of the DC-UNet model when the atrous coefficient is set to 4 are better than the case where the coefficient is 3. The recall rates are 0.56 to 0.90, indicating that the larger the expansion of the convolution kernel, the better the effect.

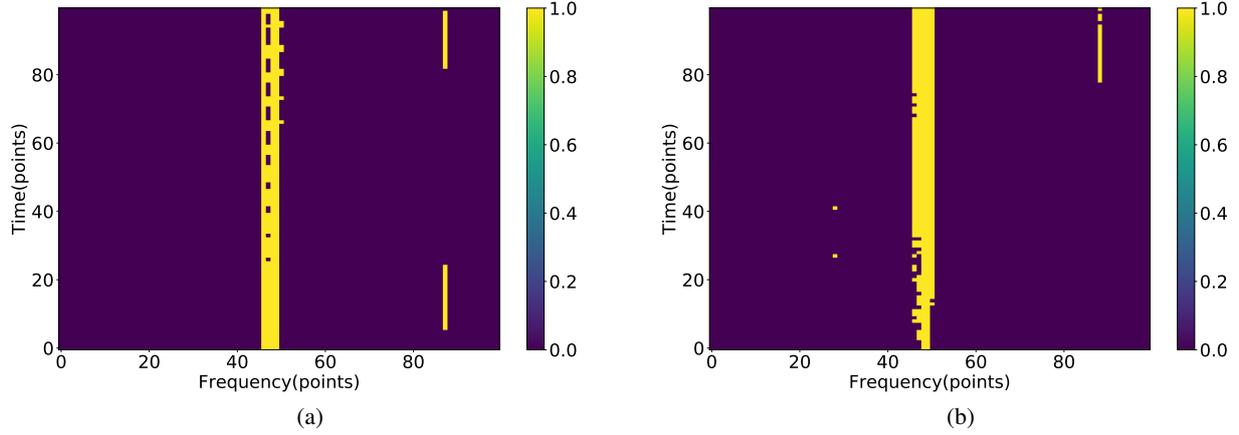


Fig. 7 Radio frequency interference mark. (a) Radio frequency interference mark output by the AC-UNet atrous_rate=7 model. (b) Radio frequency interference mark output by the U-Net model.

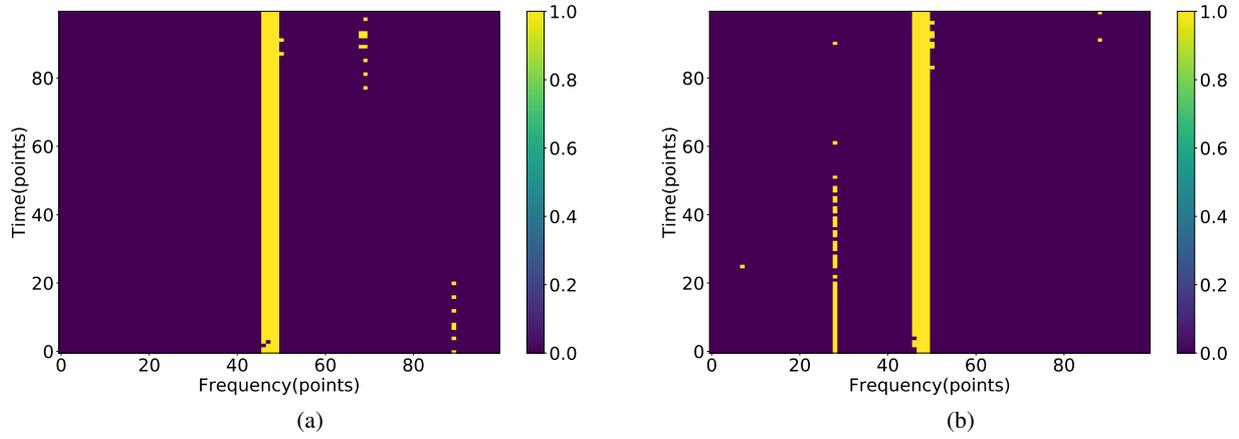


Fig. 8 Radio frequency interference mark. (a) Radio frequency interference mark output by the DC-UNet dilation_rate=4 model. (b) Radio frequency interference mark output by the AC-UNet atrous_rate=3 model.

Table 5 Experimental Results of Various Models on Data Block 512

Network Model	Classification	Recall Rate	F1-Score	AUC
U-Net	False	0.98	0.99	0.97
	True	0.96	0.83	
	avg/total	0.98	0.98	
DC-UNet dilation_rate=3	False	0.98	0.99	0.96
	True	0.93	0.84	
	avg/total	0.98	0.98	
AC-UNet atrous_rate=3	False	0.97	0.98	0.97
	True	0.96	0.74	
	avg/total	0.97	0.97	
DC-UNet dilation_rate=4	False	0.98	0.99	0.97
	True	0.95	0.79	
	avg/total	0.98	0.98	
AC-UNet atrous_rate=7	False	0.99	0.99	0.97
	True	0.96	0.86	
	avg/total	0.98	0.99	
AC-FCN	False	0.99	0.99	0.97
	True	0.96	0.81	
	avg/total	0.98	0.98	

Table 6 Experimental Results of Various Models on Data Block 1087

Network Model	Classification	Recall Rate	F1-Score	AUC
U-Net	False	0.96	0.97	0.79
	True	0.61	0.59	
	avg/total	0.94	0.94	
DC-UNet dilation_rate=3	False	0.95	0.97	0.86
	True	0.77	0.65	
	avg/total	0.94	0.94	
AC-UNet atrous_rate=3	False	0.95	0.96	0.86
	True	0.77	0.64	
	avg/total	0.93	0.94	
DC-UNet dilation_rate=4	False	0.96	0.96	0.78
	True	0.60	0.58	
	avg/total	0.93	0.94	
AC-UNet atrous_rate=7	False	0.97	0.98	0.90
	True	0.82	0.73	
	avg/total	0.95	0.96	
AC-FCN	False	0.97	0.97	0.80
	True	0.62	0.62	
	avg/total	0.94	0.94	

Then observe the various indicators values of AC-UNet atrous_rate=3 and AC-UNet atrous_rate=7. Judging

from the five test samples, the latter has a higher recall

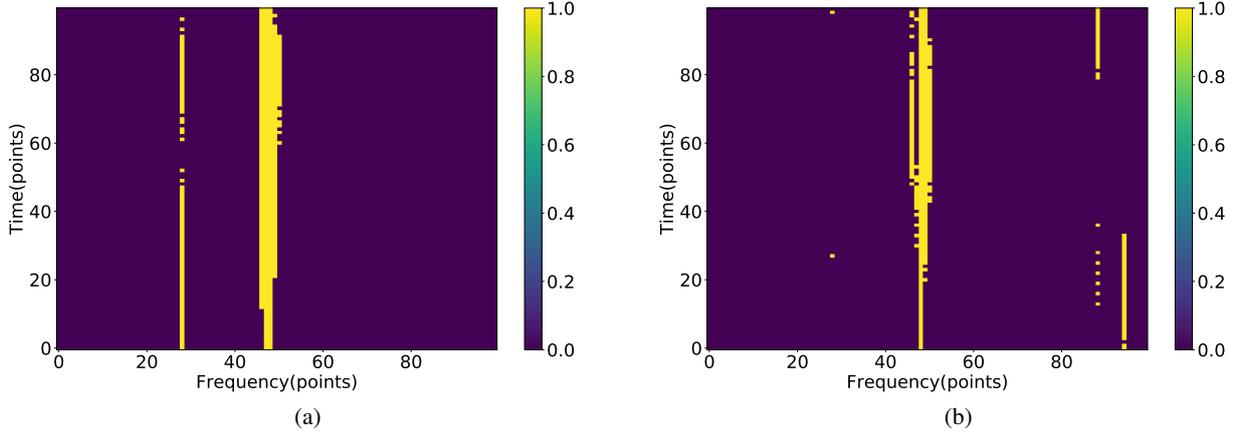


Fig. 9 Radio frequency interference mark. (a) Radio frequency interference mark output by the DC-UNet dilation_rate=3 model. (b) Radio frequency interference mark output by the AC-FCN model.

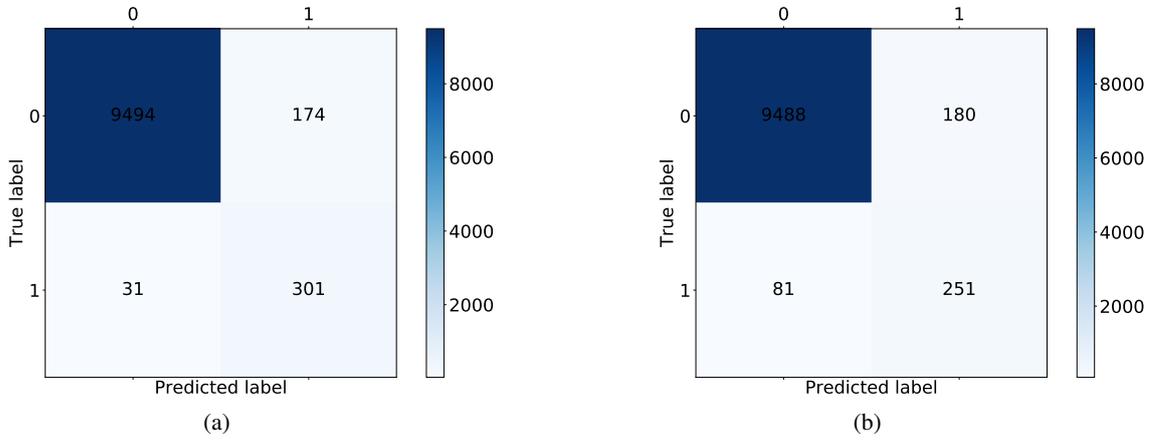


Fig. 13 Confusion matrix of different models. (a) Confusion matrix of AC-UNet model test data block 694. (b) Confusion matrix of AC-FCN model test data block 694.

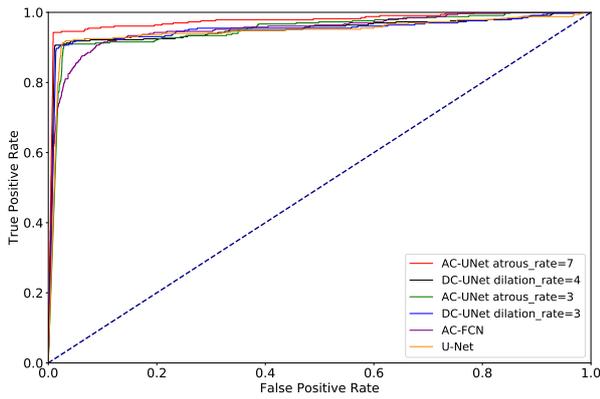


Fig. 10 ROC curve of the classification effect of each model on data block 694.

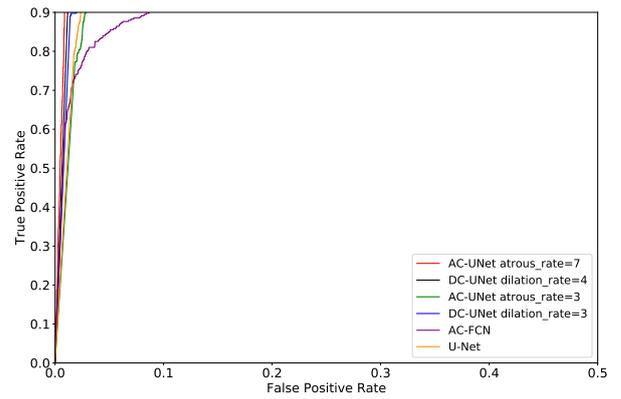


Fig. 11 ROC curve of the classification effect of each model on data block 694 (FPR ∈ [0.0, 0.5], TPR ∈ [0.0, 0.9]).

rate than the former. For the 694 data block (Table 2), the recall rate was increased from 0.90 to 0.91, the F1 score was increased from 0.68 to 0.75, and the AUC value was increased from 0.936 to 0.944, indicating that the larger the convolution kernel expansion, the better the effect.

Compared with DC-UNet dilation_rate=3 and AC-UNet atrous_rate=3, the improvement effect of DC-UNet dilation_rate=4 in various indicators is not very obvious, and even the recall rate on the 694 and 1087 data blocks

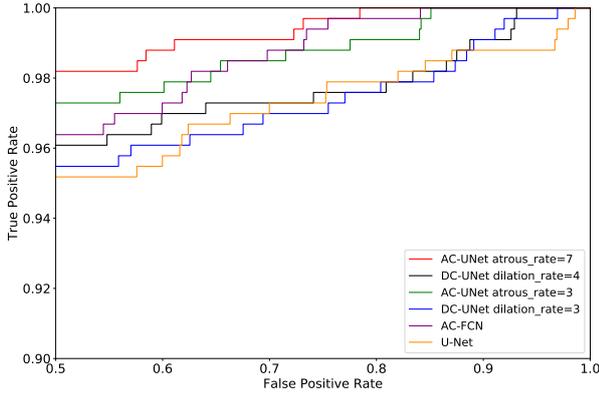


Fig. 12 ROC curve of the classification effect of each model on data block 694 ($FPR \in [0.5, 1]$, $TPR \in [0.0, 1.0]$).

has declined. Therefore, setting the expansion coefficient to 4 cannot effectively improve the segmentation effect. Finally, observe the situation of DC-UNet dilation_rate=4 and AC-UNet atrous_rate=7. Judging from these five test samples, the latter’s prediction effect is better than the former. The same is true for the 694 data block, indicating that the expansion coefficient increases and the prediction result is more accurate. Experiments verify that on this data set, the model is more suitable for implementing atrous convolution using the AtrousConvolution2D layer with an expansion coefficient of 7.

From the perspective of ROC curve, the classification effect of each model in data block 694 is shown in Figure 10. All model curves are above the 45-degree diagonal, indicating that each model is effective and can predict RFI in the image. The closer the ROC curve is to the coordinate point (0, 1), the greater the AUC value and the better the classification prediction effect for each point. In order to observe the pros and cons of each model’s prediction in more detail, the horizontal coordinate FPR is divided into $[0.0, 0.5]$ and $[0.5, 1.0]$, and the vertical coordinate TPR is divided into $[0.0, 0.9]$ and $[0.9, 1.0]$, as shown in Figure 11 and Figure 12.

It can be seen from Figure 11 that within $FPR \in [0.0, 0.5]$ and $TPR \in [0.0, 0.9]$, the red curve is relatively closer to the Y axis; it can be seen from Figure 12 that within $FPR \in [0.5, 1.0]$ and $TPR \in [0.9, 1.0]$, the red curve is relatively closer to the straight line $Y=1$. Therefore, the area between the red curve and the coordinate axis is the largest, and can correspond to the indicator AUC value.

According to the introduction of the model in Section 4, we know that our model detects the RFI in the image pixel by pixel. Therefore, our model can also detect the image with a large amount of RFI. Moreover, for images with a small amount of RFI, the model can show good performance. For a small amount of RFI, we mainly observe the overall proportion of RFI in the image.

It also can be seen from Table 1 that the proportion of RFI in the data blocks 694, 518, 510, and 512 used in our experiment is less than 5%. So, we consider it to be an image with a small amount of RFI. In addition, through the comparison of various measurement indicators and labeling result graphs in the experiment, it can be seen that the AC-UNet model can detect a small amount of RFI in the image more accurately and comprehensively than the AC-FCN model. The amount of RFI in the data blocks 694, 518, 510, and 512 is less than 500, and in the test of these data, for the 694 data block with only 332 RFI data points: as shown in Table 2, the AC-UNet atrous_rate=3 model has a higher recall rate of 0.15 than the AC-FCN model and has higher F1 score and AUC values; Regarding the confusion matrix, according to Figure 13, the AC-UNet and AC-FCN models can detect 301 and 251 data points with RFI, respectively. It shows that the method proposed in this paper can detect RFI as comprehensively and accurately as possible, and there are relatively few misjudgments.

For the 518 data block with only 401 RFI data points: the recall rate of the AC-UNet model is 0.31 higher than that of the AC-FCN, and it has a higher F1 score and AUC value. For 510 and 512 data blocks with similar recall rates, the model proposed in this paper has higher F1 score and AUC values. For the 1087 data block, the amount of RFI is relatively large, and the AC-UNet model can also get the highest recall rate, F1 score and AUC value. In summary, the AC-UNet model proposed in this paper can achieve a more robust, comprehensive and accurate detection effect on the data set.

6 CONCLUSIONS

In order to better detect radio observation images with a small amount of RFI, this paper proposes a U-Net model that can better restore image details: the AC-UNet model. At the same time, the dropout layer was added to accelerate the convergence of the model, so as to achieve the purpose of RFI detection in the observation data of the radio telescope. In order to observe the classification effect of the model on the existence and nonexistence of RFI data points, the recall rate and F1 score of the model on the existence and nonexistence of RFI are output separately, and use the AUC value to evaluate the comprehensive classification effect of the model, make the ROC curve and confusion matrix to visually display the classification effect of different models.

Through different implementation methods of atrous convolution and different expansion coefficient settings, the experimental results verify that when the expansion coefficient is 7, the AC-UNet model can play the best effect. Simultaneously set up multiple U-Net-based models for experimental comparison. The experimental results

prove that the model proposed in this paper can obtain higher recall rate, F1 score and AUC value on the test sample, and its ROC curve is closer to the upper left corner than other models. It shows that this model is better for feature extraction of the original image, can restore the original image more fully during the upsampling process, and can detect RFI in radio observation data more comprehensively and accurately.

The U-Net model was originally applied to the cell segmentation problem in the biomedical field, and it can extract many edge features in the image well. Therefore, the AC-UNet model proposed in this paper can not only be used to detect RFI, but also can be used to extract filaments in radio astronomy images. By detecting the signal and background in astronomical images, the task of detecting signals with edge characteristics in astronomical images is realized.

Acknowledgements This research was supported by the National Natural Science Foundation of China (Grant Nos. 11471045 and 41672323) and the Interdiscipline Research Funds of Beijing Normal University. We are grateful for the data provided by the National Astronomical Observatories of the Chinese Academy of Sciences. We also thank the anonymous referee for a helpful report.

References

- Akeret, J., Chang, C., Lucchi, A., & Refregier, A. 2017, *Astronomy and Computing*, 18, 35
- An, T., et al. 2017, *Acta Astronomica Sinica*, 58, 43
- Badrinarayanan, V., Kendall, A., & Cipolla, R. 2018, in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39, 2481
- Chen, L.-C., Papandreou, G., Kokkinos, I., et al. 2014, arXiv e-print, arXiv:1412.7062
- Chen, L.-C., Papandreou, G., Kokkinos, I., et al. 2017a, in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40, 834
- Chen, L.-C., Papandreou, G., Schroff, F., et al. 2017b, arXiv e-print, arXiv:1706.05587
- Chen, X., Wu, F., Shi, H., et al. 2015, *IAU General Assembly*, 22, 2252187
- Czech, D., Mishra, A., & Inggis, M. 2017, *Radio Science*, 52, 841
- Deng, L., & Yu, D. 2014, in *Foundations and Trends in Signal Processing*, 7, 197
- Fayjie, A. R., Hossain, S., Oualid, D., et al. 2018, in 2018 15th International Conference on Ubiquitous Robots (UR), 896
- Fridman, P., & Baan, W. 2001, *A&A*, 378, 327
- Greenspan, H., van Ginneken, B., & Summers, R. M. 2016, in *IEEE Transactions on Medical Imaging*, 35, 1153
- He, K., Gkioxari, G., Dollár, P., et al. 2017, in *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 42, 386
- He, K., Zhang, X., Ren, S., & Sun, J. 2016, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770
- Heaton, J. B., Polson, N. G., & Witte, J. H. 2016, arXiv e-print, arXiv:1602.06561
- Kingma, D., & Ba, J. 2014, *Computer Science*, <https://arxiv.org/pdf/1412.6980.pdf>
- Kocz, J., Briggs, F. H., & Reynolds, J. 2010, *AJ*, 140, 2086
- Lahtinen, J., Uusitalo, J., Ruokokoski, T., et al. 2017, in *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 10, 3087
- Li, J., Zuo, S., Wu, F., et al. 2020, *Science China Physics, Mechanics, and Astronomy*, 63, 129862
- Liu, W., Anguelov, D., Erhan, D., et al. 2016, in *Lecture Notes in Computer Science*, eds. Leibe, B., Matas, J., Sebe, N., & Welling, M. (Computer Vision) ECCV 2016, 9905, 21 (Cham: Springer)
- Long, J., Shelhamer, E., & Darrell, T. 2015, in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1, 3431
- Offringa, A. R., de Bruyn, A. G., Zaroubi, S., et al. 2013, *A&A*, 549, A11
- Offringa, A. R., de Bruyn, A. G., Biehl, M., et al. 2010, *MNRAS*, 405, 155
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. 2016, in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, 779
- Ren, S., He, K., Girshick, R., & Sun, J. 2015, in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39, 1137
- Ronneberger, O., Fischer, P., & Brox, T. 2015, *Medical Image Computing and Computer-Assisted Intervention*, eds. Navab, N., Hornegger, J., Wells, W., & Frangi, A., MICCAI 2015, 9351, 234 (Cham: Springer)
- Sainath, T. N., Kingsbury, B., Saon, G., et al. 2015, *Neural Networks*, 64, 39
- Schmidhuber, J. 2015, *Neural Networks*, 61, 85
- Simonyan, K., & Zisserman, A. 2014, arXiv e-print, arXiv:1409.1556
- Wu, F., Wang, Y., Zhang, J., et al. 2014, in 2014 XXXIth URSI General Assembly and Scientific Symposium (URSI GASS), IEEE, 1
- Xia, X., Xu, C., & Nan, B. 2017, in 2017 2nd International Conference on Image, Vision and Computing (ICIVC), 783
- Zhu, S., Wang, Z., Wang, M., & Dong, L. 2017, *Astronomical Research & Technology*, 014, 297 (in Chinese)