DSC based Dual-Resunet for radio frequency interference identification

Yan-Jun Zhang[1], Yan-Zuo Li[2], Jun Cheng[3,4]⋆ and Yi-Hua Yan[3]

[1] School of Cyberspace Science and Technology, Beijing Institute of Technology, Beijing 100081, China; *zhangyj@bit.edu.cn*
[2] School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China; *lyz@bit.edu.cn*
[3] CAS Key Laboratory of Solar Activity, National Astronomical Observatories, Beijing 100101, China; *chengjun@nao.cas.cn,yyh@nao.cas.cn*
[4] State Key Laboratory of Space Weather, Chinese Academy of Sciences, Beijing 100864, China

Abstract Radio frequency interference (RFI) will pollute the weak astronomical signals received by radio telescopes, which in return will seriously affect the time-domain astronomical observation and research. In this paper, we use a deep learning method to identify RFI in frequency spectrum data, and propose a neural network based on Unet that combines the principles of depthwise separable convolution and residual, named DSC Based Dual-Resunet. Compared with the existing Unet network, DSC Based Dual-Resunet performs better in terms of accuracy, F1 score, and MIoU, and is also better in terms of computation cost where the model size and parameter amount are 12.5% of Unet and the amount of computation is 38% of Unet. The experimental results show that the proposed network is a high-performance and lightweight network, and it is hopeful to be applied to RFI identification of radio telescopes on a large scale.

Key words: techniques: deep learning and image processing — radio frequency interference — telescopes — Sun: radio radiation

1 INTRODUCTION

With the increasing sensitivity of radio telescopes, large-scale visits, and the use of low-power artificial broadband signals, the signals received by radio telescopes produce a lot of interference, which seriously affects the research of astronomical observation data. To reduce radio frequency interference (RFI), radio telescope sites are usually located in radio-quiet zones, and a series of hardware optimization designs, such as improvement of grounding shield and band-pass filter, are needed to avoid mutual inductance coupling induction, common impedance induction, field-induced induction and other main paths of RFI generation. In addition, post-processed of the observed data is required to remove RFI.

In the early days, astronomers manually marked RFI by visual inspection, which can be done in standard astronomical imaging software. However, with the emergence of higher spectral and temporal resolution, astronomy has ushered in a new stage of development, i.e., the full-band-large-sample-massive information period, where the amount of data is getting larger and larger, and

manual marking suffers from many problems such as time-consuming and human interference. The basic idea of automatic RFI identification is to label RFI accurately through statistical analysis of data, i.e., eliminating RFI by the difference of morphological characteristics between RFI signal and solar radio observation signal in time-frequency two-dimensional map. The current post-processing elimination methods can be divided into three categories: 1) Threshold-based methods, such as cumulative sum method and combined threshold sum method. These methods define RFI as pixels in spectrum that exceed certain thresholds. This kind of algorithm is simple and efficient. Astronomical equipment such as LOFAR (Cendes et al. 2014) uses this method to process RFI. The biggest problem of this method is that it is difficult to determine the threshold according to RFI source and observation signal, especially for time-varying celestial signal, where the selection of threshold is particularly critical and sometimes the weak instantaneous signal may be identified as RFI and removed; 2) Wavelet-based methods. Since wavelet transform has the characteristics of time-frequency localization and variable resolution, the distribution characteristics of signal and noise are different in wavelet domain, and Mészárosová et al. (1999)

proposed to use wavelet analysis to study two types of solar radio bursts. However, the difference in time-frequency characteristics of RFI from different sources makes it very complex and difficult to model all types of RFI signals. The method designed for a particular event is not universally applicable to other cases. 3) Machine learning-based methods. At present, more and more researchers began to use this method to solve the RFI problems in the field of astronomy. This is because machine learning has made remarkable achievements in many fields, and in the field of astronomy, there is enough data for training. Initially, researchers tried to identify the characteristics of RFI through linear methods, such as singular vector decomposition (SVD) (Offringa et al. 2010) or principal component analysis (PCA) (Zhao et al. 2013). If the RFI pattern exhibits repetitive patterns in time and frequency, these methods work well, but they cannot handle more random signals, such as those caused by irregular satellites. Later, they use traditional supervised machine learning techniques. In literature (Wolfaardt 2016), RFI signals are clustered based on K-nearest neighbor and Gaussian mixture model to realize RFI labeling. Bethapudi & Desai (2018) analyzed and evaluated the performance of four different supervised learning algorithms: an Artificial Neural Network Multi-Layer Perceptron (ANN MLP), Adaboost, Gradient Boosting Classifier (GBC), and XGBoost, for the separation of pulsars from RFI and other sources of noise. However, for these methods to achieve a sufficient classification accuracy, a careful feature selection process has to be performed prior to the application. With the emergence of deep learning, algorithm can automatically learn features from a large number of image data using network structures of different scales and various learning rules without prior knowledge and feature selection. Some researchers tend to use deep learning methods for RFI recognition research. Daniel (Czech et al. 2018) demonstrated an approach to classifying the sources of transient RFI (in time domain data) that use CNNs and LSTMs. Akeret et al. (2017) employed a special type of Convolutional Neural Network, the U-Net, that enables the classification of clean signal and RFI signatures in 2D time-ordered data. Dai et al. (2019) also used the Unet to detect the RFI signals on Tianlai data.

It is not difficult to see from the above, the research on using deep learning to identify RFI is relatively few and not perfect. Akeret et al. (2017) mentioned above used the original Unet (Ronneberger et al. 2015) for medical image processing. In recent years, scholars have made many improvements based on the original Unet. However, these research results have not been used in RFI identification. Zhang et al. (2020) mainly discussed the influence of depth of Unet model on network performance. Guan et al. (2018) improved the extraction of features by

replacing the original two ordinary convolutions with a densely connected structure, which directly connects all layers while ensuring the maximum information transfer between layers in the network, reducing disappearance of gradients and making the network deeper. At the same time, it strengthens the feature transfer between layers and the feature is used more effectively. Liu et al. (2020) used another method, residual connection, to solve the problem of degradation and gradient disappearance of deep networks, which can make network converge faster and get better results while deepening the network. Hu et al. (2019) considered from the perspective of perceptual field size of the neural network. They firstly used convolutional kernels of different sizes to extract feature and then feature fusion was performed. This increases the width of network and the adaptability to feature size, and achieves good results. Gong et al. (2020) also considered the influence of the receptive field, and since computation corresponding to the increases of convolution kernel grows geometrically, they used expansion convolution to change the size of the receptive field and convolution kernels with different expansion coefficients in each convolution layer. In addition, some scholars have added attention mechanisms to network to improve network performance. Huang et al. (2017) proposed a reverse attention structure that generates a mask for each class to amplify the reverse class response and learn the content that is not related to the region of interest. Alternatively, a feature pyramid network (Li et al. 2018) is introduced to generate attention signals from different pyramid scales and performs a global average pooling operation to provide global context as a guide to low-level features to compute category location details. Chen et al. (2021) introduced SE module into the network which is a novel feature recalibrating strategy. Specifically, the importance of each feature channel is automatically obtained through learning and then valuable features are promoted and features that are not useful for the current task are suppressed according to this importance.

Based on the experience of the above literature, and according to the characteristics of RFI in the spectrum, the authors proposed DSC Based Dual-Resunet in this paper. DSC Based Dual-Resunet is based on the architecture of Unet network, in which a dual Unet structure is designed using Unets stacking, and a residual block is designed to replace the standard convolution for better extraction of image features. At the same time, depthwise separable convolution is introduced to make the network more lightweight and improve computational efficiency. In addition, according to the unique characteristics of time-frequency two-dimensional graph containing RFI, the weight coefficient of jump connection is proposed to better extract low-level features. For the problem that the
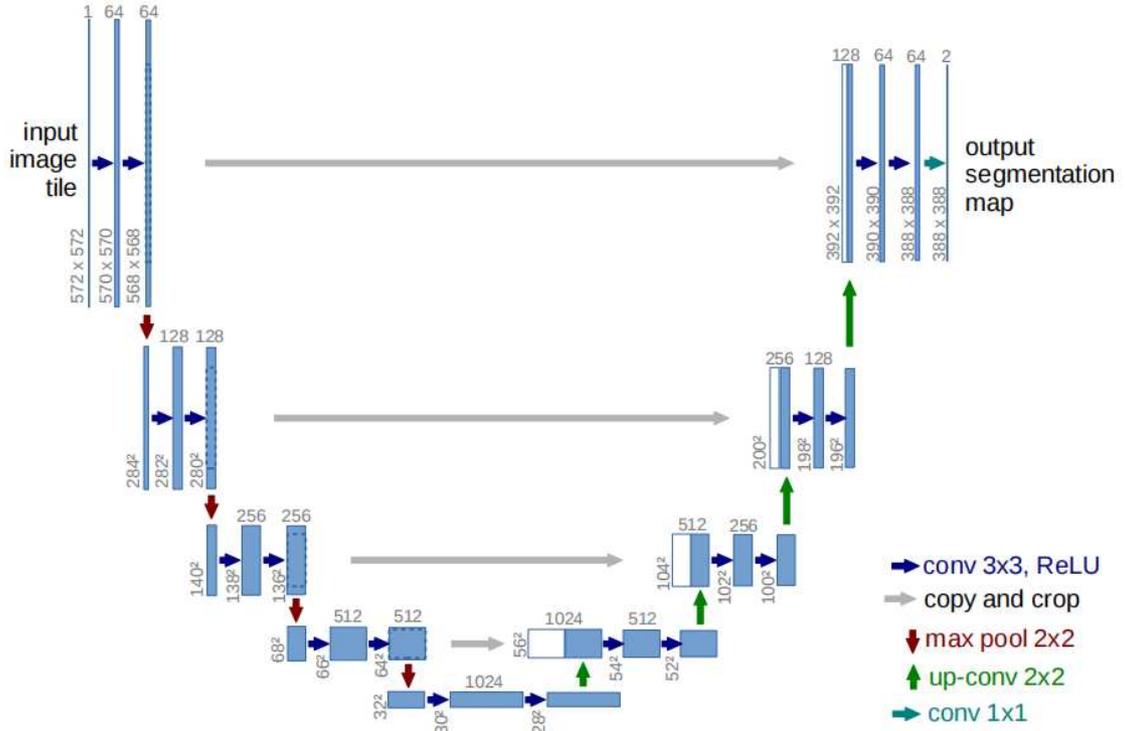
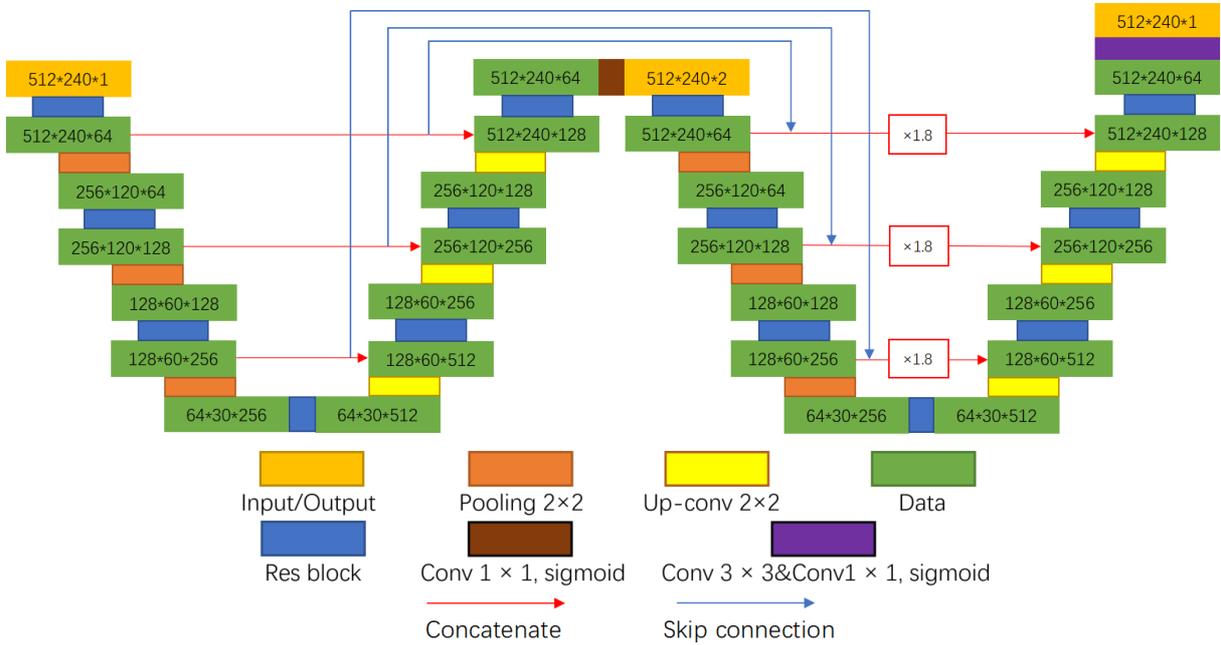**Fig. 1** Unet network structure.

**Fig. 2** DSC Based Dual-Resunet network Structure.

proportion of positive and negative samples in data varies greatly, dice loss is used to avoid training failure.

This paper is organized as follows: in Section 2, the structure of the proposed neural network model and its advantages are introduced. Section 3 trains and tests the proposed network model, and the model is evaluated with performance metrics, compared with Unet, and analyzed. Finally, summaries and conclusions are given in Section 4.

2 NETWORK MODEL

In this section, the Unet model is introduced firstly, and then the proposed network is described in detail.
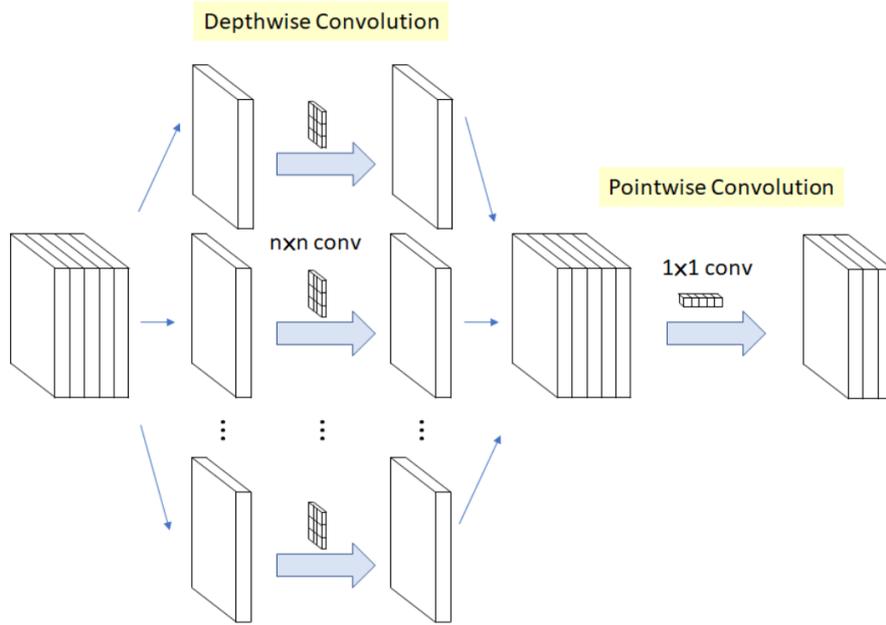
Fig. 3 Schematic diagram of the depthwise separable convolution principle.

The related principles such as residual connection and depthwise separable convolution used in the network are introduced. This network structure is named DSC Based Dual-Resunet which is especially proposed according to the characteristics of RFI that need to be identified. The network can better extract the characteristics of RFI in spectrum image and learn them, and output accurate recognition results.

2.1 Unet Network Structure

Unet is a semantic segmentation model inspired by fully convolutional networks (FCN) and designed for medical images, which can learn image features using a small amount of data, and it is a very robust model for edge extraction and plays a significant role in biomedical image segmentation. The structure of Unet is shown in Figure 1.

The whole network structure is roughly divided into a contraction path and an expansion path. Such a structure is also called the encoder-decoder structure in some literature. Because the whole network structure is similar to the capital letter U, it is named Unet. The contraction path is composed of convolution and pooling components of the traditional convolution neural network. The size of feature map is reduced by pooling, and feature dimension is expanded by convolution to extract high-level features. The expansion path uses 2×2 de-convolution instead of pooling in the contraction path. After each de-convolution, image size is doubled and concatenated with the original feature map of same scale layer. De-convolutions also make the edge information of segmentation image more refined. In the last layer, the required classification results

are obtained by a convolution with a kernel size of 1×1 and activation function. There are 23 convolution layers in the whole Unet structure.

Unet is very different from other common segmentation networks: Unet uses a completely different feature fusion method, concatenating. That is, the fusion of new feature map in expansion path and feature map from contraction path. Unet concatenates features together in the channel dimension to form thicker feature maps, and the same stage uses jump connection to connect. This ensures that the final recovered feature map incorporates more low-level features, and also ensures the fusion of features of different scales, thus allowing multi-scale prediction and deep supervision. This is one of the important reasons why Unet is superior to other segmentation network models. In contrast, the corresponding points are summed in FCN and the feature maps do not thicken.

2.2 DSC Based Dual-Resunet Network Structure

Unet is mainly used for medical image segmentation, and spectrum of post-processing data of a radio telescope is similar to but not identical to a medical image, so the network structure in Figure 2 is designed according to its characteristics.

As shown in Figure 2, the network model consists of two U-shaped network structures, each contains four scale layers. The two U-shaped structures are connected by a convolution layer and three jump connections. In the structure, each scale layer contains a residual block to extract image features, and each residual block contains three depthwise separable convolutions. The scale
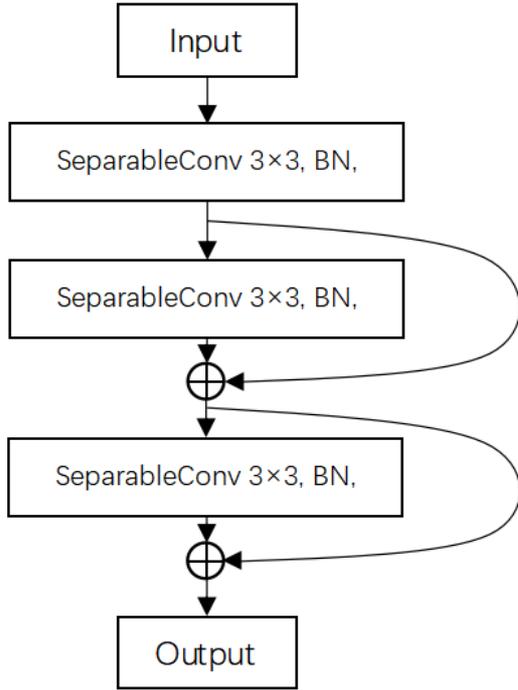
Fig. 4 Residual block structure.

Fig. 5 Four prediction results of the binary classification problem.

Fig. 6 Fitting curve of relationship between weight coefficient of jump connection and network performance.

transformation between different scale layers is completed by de-convolution or pooling. The difference between two U-shaped structures is that a weight coefficient is added to the jump connection in the latter structure. The network ends with a sigmoid function to output results. The specific structure and functions of these components are introduced in the following sections.

2.2.1 Depthwise separable convolution

Depthwise separable convolution first appeared in a Ph.D. thesis entitled "Rigid-motion scattering for image classification". However, what makes everyone familiar with DSC are two well-known models, Xception (Chollet 2016) and MobileNet (Howard et al. 2017). They are two important achievements from Google team in the same period.

The detailed structure of DSC is shown in Figure 3. DSC is composed of two parts, namely, Depthwise Convolution and Pointwise Convolution. The calculation of Depthwise Convolution is simple, it uses a convolution kernel for each channel of input feature map and then concatenates outputs of each convolution to get its final output. In Depthwise Convolution, only one convolution kernel is used for each channel, so the number of output channels is the same as the number of input channels.

Pointwise Convolution uses a specified number of 1×1 convolution kernels to convolute the output of Depthwise Convolution, which plays two roles in DSC. The first is to allow DSC to change the number of output channels freely; the second role is to fuse feature map output by Depthwise Convolution.
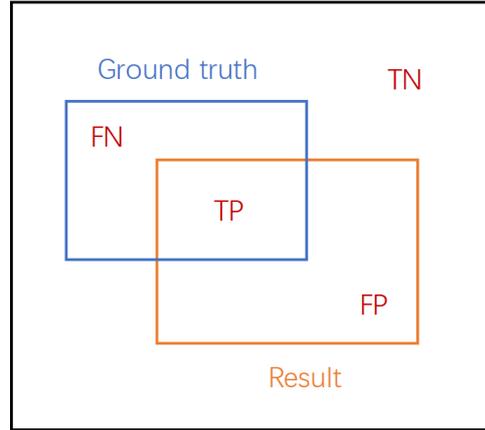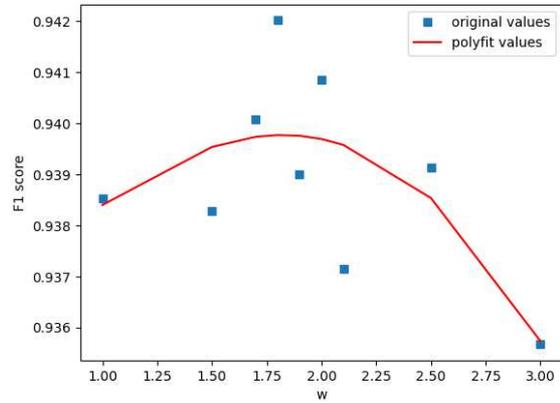
DSC is a substitute for ordinary convolution, in which the number of convolution kernels is equal to the number of input channels multiplied by the number of output channels, and the size of convolution kernels is $n\times n$. But DSC only needs the same amount of 1×1 convolution kernels and $n\times n$ convolution kernels which is equal to the number of input channels. Obviously, its computational complexity is much lower than that of ordinary convolution, especially when n is very large. So its most significant advantage is high computational efficiency and it is often used to build lightweight models, but the efficiency of DSC is at the cost of reducing network accuracy. Designing high-efficiency and high-precision DSC models is one of the hot research directions. In the proposed model, all convolutions use DSC to get higher computational efficiency.

2.2.2 *Residual block and residual connection*

In the original Unet model, image features are extracted by two convolutions in each scale layer. In the proposed model, the residual blocks are used to extract features, and the structure of residual block is shown in Figure 4.

As shown in Figure 4, a residual block contains three convolution layers and two jump connections. The jump connection here is the residual connection. The idea of residual connection appeared in the early traditional neural network. Srivastava et al. (2015) proposed the structure of residual connection, then He et al. (2015) proposed the residual network and promoted this structure. This structure is also added to the proposed network, which can improve the gradient dissipation problem in back-propagation process and make network structure become deeper, while preventing the degeneration of neural networks, and accelerating convergence process.

In this structure, each residual block contains three convolution layers but only two residual connections. Setting the number of convolution layers to 3 is the result of considering both network performance and efficiency. The whole network contains 14 residual blocks, the number of convolution layers in residual blocks has a significant impact on the whole network. Reducing the number of convolution layers will reduce network performance. Increasing the number of convolution layers will lead to an oversized network and may cause problems such as memory overflow. According to experiments, it is verified that three convolution layers are the best choice. In traditional residual network, the number of residual connections is always the same as the number of convolution layers. However, this network only uses two residual connections because the first convolution in the block will change the feature dimension. So the two ends of the first convolution cannot be directly connected. If the structure needs three residual connections, it also needs an additional convolution layer to make the feature dimensions of both ends of the residual connection consistent. This will increase the amount of computation, and it is verified that three residual connections cannot improve the performance of network.

2.2.3 *Dice loss function*

The most common loss function used by the neural networks to solve binary classification problems is a binary cross-entropy loss function, which is also used by the original author of Unet. But for this paper, the proportion of RFI is less than 5% in a time-frequency two-dimensional image, which means that positive and negative sample proportion of the binary classification problem is very different. In this case, the binary cross-entropy loss function is more likely to focus on samples with a large
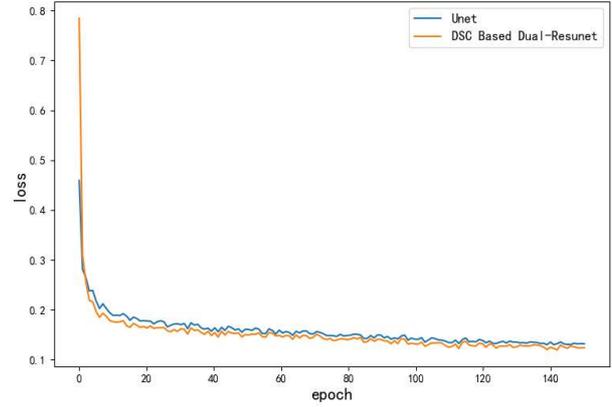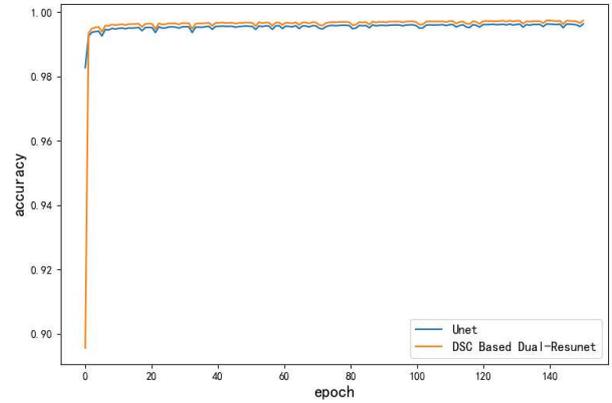

Fig. 7 Training loss function curve.

Fig. 8 Training accuracy curve.

sample ratio, and the trained network model has a poor resolution for samples with a small sample ratio. The cross-entropy function with weight, dice loss, focal loss, and other functions can be used to solve this problem. After verification, the best dice loss is used in this paper.

Firstly, the dice coefficient is an ensemble similarity measurement function, which is usually used to evaluate the similarity of two samples, and the larger the value is, the more similar it is. The value range is [0,1] and the dice coefficient is expressed as:

$$s = \frac{2|X \bigcap Y|}{|X| + |Y|}, \qquad (1)$$

where $|X \bigcap Y|$ is the intersection between X and Y, $|X|$ and $|Y|$ indicate the number of elements of X and Y, respectively, where the coefficient of molecule is 2, because the denominator repeatedly calculates the common elements between X and Y. So the dice loss is expressed as:

$$\text{dice} = 1 - s. \qquad (2)$$

For binary classification problems, the general predictive results are as follows: TP: true positive, TN: true negative, FP: false positive, FN: false negative. They are schematically shown in Figure 5.
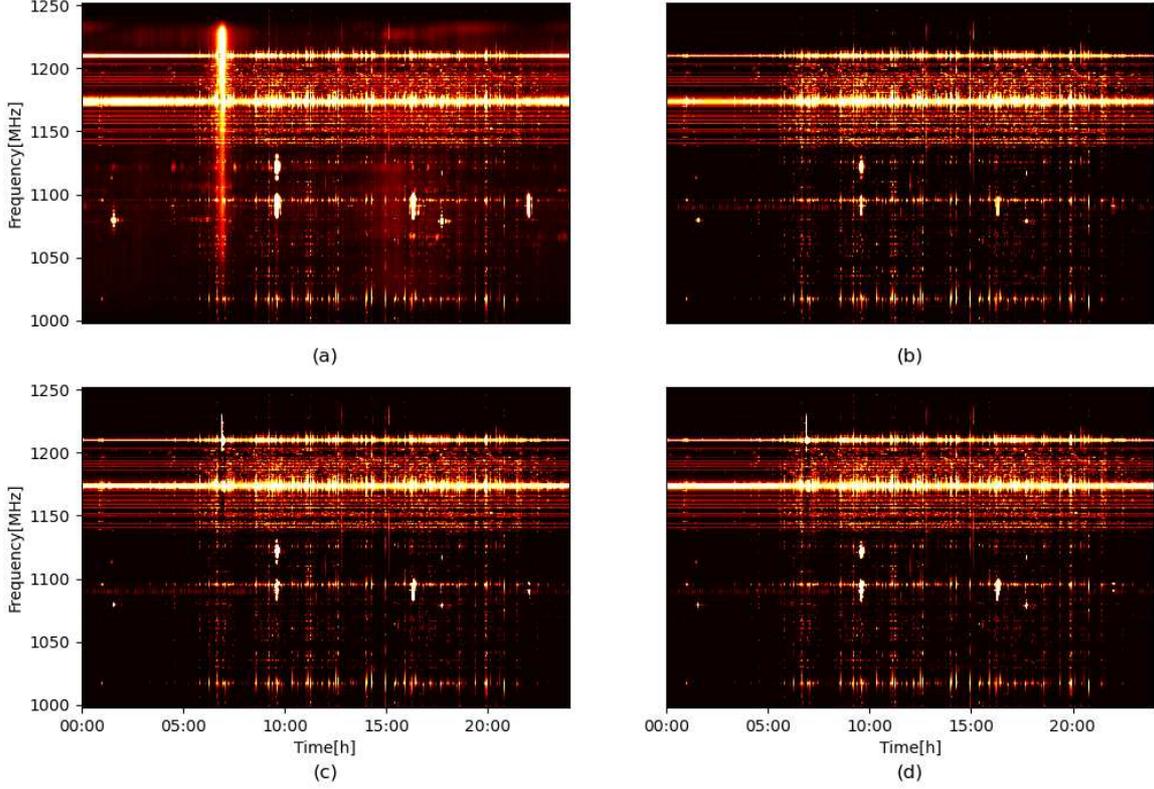
Fig. 9 Network classification results. (a) is the observed TOD from the Bleien Observatory with RFI, (b) is the RFI mask obtained from SEEK's SumThreshold , (c) and (d) are the RFI mask from the proposed network and Unet, respectively.

The dice coefficient can also be written as follows:

$$s = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}}. \tag{3}$$

The precision, recall rate, and F1 are expressed as:

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{4}$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{5}$$

$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} = s. \tag{6}$$

It can be seen that the dice coefficient is equivalent to F1 score. Dice coefficient is used to calculate similarity between $|X|$ and $|Y|$, but essentially contains two metrics, precision and recall. Therefore, dice loss is an optimization of F1 score, and F1 score is also used as one of the reference metrics in the subsequent evaluation of network performance.

Dice loss can be used to train network model with a stable convergence process and has a better performance compared with other loss functions, which can solve the problem caused by extremely unbalanced ratio of positive and negative samples.

2.2.4 Weight coefficient of jump connection

In the proposed network structure, a weight adjustment coefficient is added to the jump connection of the second U-shaped structure, which is named weight coefficient of the jump connection. In essence, the double Unet structure increases the number of convolutional layers compared to a single Unet, and the whole network becomes deeper. It is known that shallow network extracts low-level features of image, and deep network extracts high-level features, and as the number of network layers increases, the influence of low-level features gradually weakens, which is not conducive to the extraction of global features of image. However, RFI in time-frequency two-dimensional diagram is scattered, and its features are relatively simple which is low-level features, so low-level features are more important than high-level features. Moreover, the proposed network structure is deeper, so the jump connection weight coefficient is introduced to improve influence of low-level features. This makes low-level features better transmitted to the back-end of network, the larger the weight is, and the greater the influence of front-end features have on back-end.

To determine the optimal value of weight coefficient, weight coefficients of different values is used to train the proposed network, and F1 score is used as evaluation
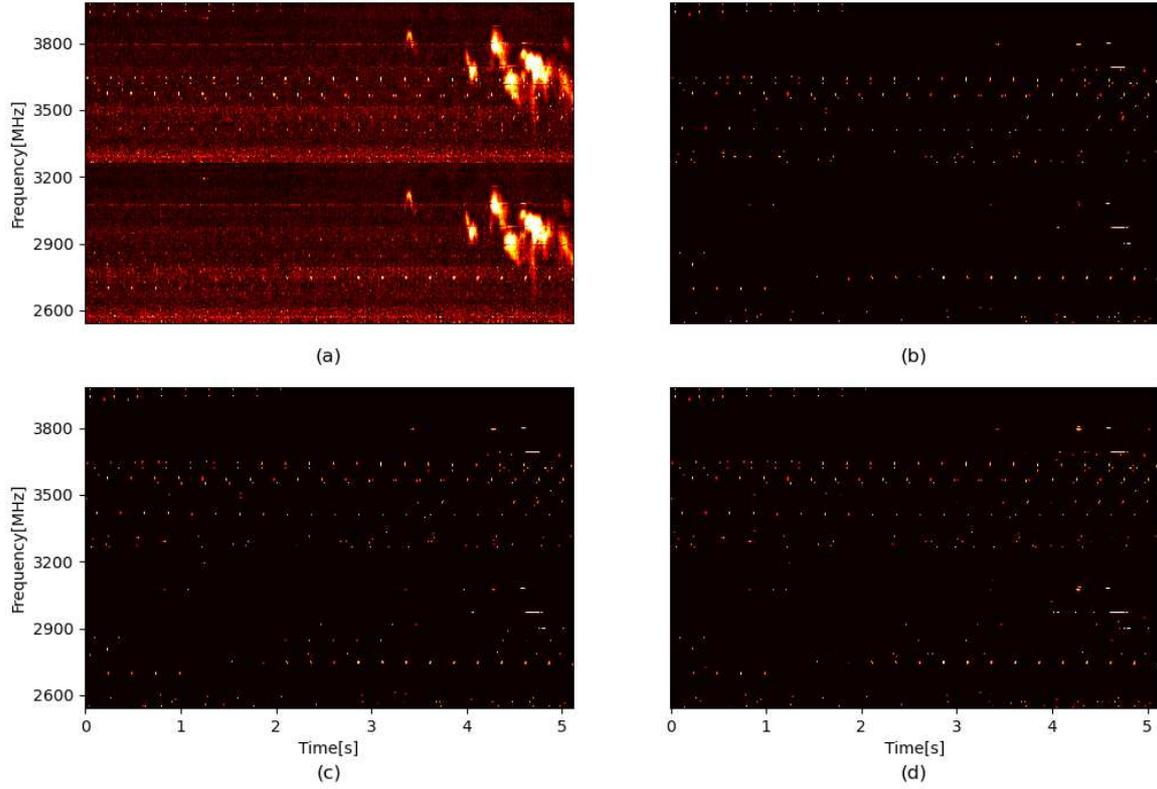
Fig. 10 Network classification results. (a) is the observed TOD from SBRS with RFI, (b) is the classification results of SEEK's SumThreshold , (c) and (d) are the classification results of the proposed network and Unet, respectively.

metric to curve fit the experimental results as shown in Figure 6, The results are judged according to the curve fitting and combined with original value before fitting, and finally w=1.8 is selected and achieved better results in the experiment.

3 EXPERIMENT AND RESULTS ANALYSIS

3.1 Experimental Data

In order to compare the performance of the two networks, firstly, we used the same data set as Akeret et al. (2017), which is taken at Bleien Observatory. Then we used the solar radio spectrum data observed by Chinese Solar Broadband Radio Spectrometer (SBRS) and specially selected some data containing burst. The two groups of data sets were processed and tested by the same method. The data sets was also labeled by SumThreshold (Offringa et al. 2010) algorithm, which is a variable threshold method for improving the classification performance in which the sum of one or more samples is used as the threshold criterion, and the basic principle to judge whether a pixel exceeds the threshold is that if A and B do not exceed the single sample threshold X_1, but the average value of A and B exceeds the slightly lower threshold X_2, they can still be marked, if not, they can be combined with the third neighbor C and threshold at

X_3. The more samples combined, the lower the sample threshold is, and so on. The number of samples selected for each judgment is called combination number.

The threshold for different combination number i is determined by two parameters, X_i (threshold of single sample) and ρ. X_i is represented by the following formula:

$$X_i = \frac{X_1}{\rho^{\log_2 i}} . \tag{7}$$

According to the experience, the value of $\rho = 1.5$ is suitable for most cases.

The specific implementation process is as follows: first, the image is smoothed by one-dimensional Gaussian filtering in two directions, and then the residual image which contains systematic noise and RFI is obtained by subtracting the smoothed data from the original data. The residual image is then used to label RFI through the combined threshold method mentioned above, so as to avoid marking astronomical sources with high amplitude visibility.

3.2 Performance Evaluation Metrics

In this paper, the proposed neural network model is to identify RFI, which is essentially an image segmentation problem and also a pixel-specific binary classification problem. The evaluation metric for the classification
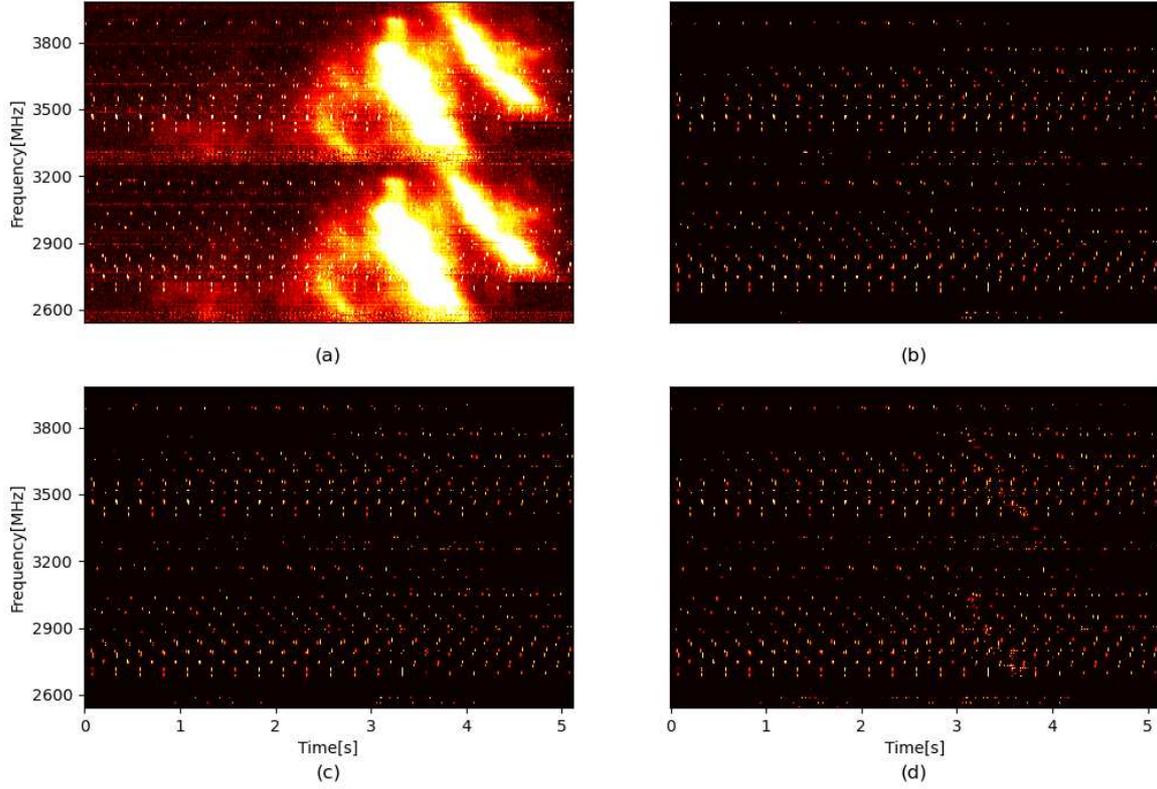
Fig. 11 Network classification results. (a) is the observed TOD from SBRS with RFI, (b) is the classification results of SEEK's SumThreshold, (c) and (d) are the classification results of the proposed network and Unet, respectively.

problem is firstly accuracy, but accuracy cannot fully and objectively reflect the learning ability of neural networks when the proportion difference between positive and negative samples is too large, and the commonly used evaluation metrics in the field of image segmentation are F1 score and MIoU.

To comprehensively evaluate the classification performance of models, it is necessary to check precision rate and recall rate at the same time. Unfortunately, precision rate and recall rate are often negatively correlated. In other words, increasing precision usually reduces recall value. F1 score is a metric used to measure the accuracy of binary classification or multitask classification model in statistics and takes into account both precision and recall of classification model. It can be regarded as a weighted average of precision and recall, with the maximum value of 1 and the minimum value of 0. A larger value means better classification performance.

Precision refers to the proportion of samples with a predicted value of 1 and a true value of 1 in all samples with a predicted value of 1. Recall refers to the proportion of samples with a predicted value of 1 and a true value of 1 in all samples with a true value of 1. The calculation methods of precision, recall and F1 score are given in Equations (4), (5), (6) respectively.

In the field of image segmentation, MIoU value is also an important metric to measure the accuracy of image segmentation. MIoU can be interpreted as average cross-union ratio, which means average of IoU value of each category. It is a global-based evaluation. Among them:

$$\mathrm{IoU} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP} + \mathrm{FN}}, \tag{8}$$

$$\begin{aligned}
\mathrm{MIoU} &= \frac{1}{k+1} \sum_{i=0}^{k} \frac{p_{ii}}{\sum_{j=0}^{k} p_{ij} + \sum_{j=0}^{k} p_{ji} - p_{ii}} \\
&= \frac{1}{k+1} \sum_{i=0}^{k} \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP} + \mathrm{FN}}.
\end{aligned} \tag{9}$$

To compare classification performance of neural networks, accuracy, F1 score and MIoU are used as the evaluation criteria, but for a network model, its computation and model size are also important metrics, both of which correspond to time complexity and space complexity, respectively. The condition for implementation of algorithm model is to meet needs of business scenarios such as memory occupation and computing speed while ensuring the performance of algorithm. It is very important for model deployment of mobile terminal, so the size of model and the amount of computation have an important impact on the practical application of model.

Table 1 Comparison of Performance and Computing Cost of Network Model

Model	Accuracy	F1	MIoU	FLOPs(G)	Params(M)	Model Size(MB)
Unet	0.98526	0.93267	0.93499	192.9	31.1	372.9
Proposed Network	0.98557	0.93958	0.94156	73.5	3.9	47.7

These two metrics and the number of parameters with similar effect to the model size are also added to the evaluation metrics.

3.3 Training and Testing

This section uses the dataset in Section 3.1 to train and verify networks. 60% of the dataset is taken as training samples and the rest as test samples. The size of each sample is 240×512. In the pre-training, it is found that the loss of the training model tends to be stable after 150 epochs, so all models are set to 150 epochs for each training. The initial weight value of each layer is random. Dice loss is used as the loss function, Adam is used as the optimizer, and the learning rate is set to 0.0001. In the process of training, methods of data enhancement such as translation and rotation are also used to expand diversity of the data set and improve the learning efficiency of neural networks.

The training and testing are implemented in the software environment of Ubuntu 18.04 + python 3.6 + TensorFlow 1.15.0 with a hardware platform of an Nvidia GTX3080 graphics card. The Unet and the proposed network are trained and tested under the same conditions. The results of experiment are averaged and analyzed.

3.4 Experimental Results and Analysis

After training and testing the neural networks under the above conditions, a series of results are obtained. The experimental results are summarized as mean values and used for qualitative and quantitative evaluation of the neural networks.

For training process, the training loss curve and the training accuracy of each network are shown in Figures 7 and 8.

According to Figures 7 and 8, it can be seen that DSC Based Dual-Resunet converges faster during training process, due to the design of the residual block, which makes the network deeper and takes advantage of the residual connection. While Unet does not use a residual connection, so the performance of Unet is not as good as the proposed network. At the same time, the performance of the proposed network is also better for loss and accuracy value in training.

Figure 9 shows examples for classification performance comparison, in which the subgraphs from top to bottom are the observed TOD from the Bleien Observatory with RFI, the RFI mask obtained from SEEKs SumThreshold the RFI mask from the proposed network, and the RFI mask from Unet, respectively. It can be seen that both Unet and the proposed network can identify most RFI accurately, and only a small number of RFI are not marked. But the proposed network is better than Unet. The prediction result of the proposed network visually is closer to the truth diagram, and the data statistical result is also like this.

After multiple rounds of tests, all the test results are averaged and summarized in the Table 1. Among them, floating-point operations (FLOPs), which represent the amount of computation, is the multiplication and addition required for a 240×512 sample to be processed once by a neural network.

As can be seen from the Table 1, DSC Based Dual-Resunet has better network performance in terms of accuracy, F1 score, and MIoU. As for the computation cost of the models, DSC Based Dual-Resunet is far superior to Unet in terms of model size and the amounts of parameters and computation, due to the introduction of depthwise separable convolution. The model size and the number of parameters are 12.5% of Unet, and the amount of computation is 38% of Unet. It can be seen that the model size, the amounts of parameters, and computation are greatly reduced. The amount of computation determines the consumption of computing resources and training time, and the model size determines the memory occupied by model in application. For deep learning network models, lightweight is one of the key directions of research, that is, reducing size and parameters of model as much as possible. Because the fewer the amount of computation, the faster the computation speed, and lightweight networks can be more easily applied in mobile hardware. The network proposed in this paper is a lightweight model, which has fewer parameters and smaller model size than other models, and the computing speed is greatly improved. In conclusion, compared with Unet, DSC Based Dual-Resunet has better performance, lower time complexity and space complexity, which means that the model can save more computing resources and is more efficient.

Figures 10 and 11 also show examples for classification performance comparison, in which the subgraphs from top to bottom are the observed TOD from SBRS with RFI to be predicted, the RFI mask obtained from SEEK's SumThreshold, the RFI mask from the proposed network, and the RFI mask from Unet, respectively. The

results showed that the proposed network can identify most RFI in application and distinguish between burst and RFI signals. The prediction results of the proposed network are significantly better than Unet in burst areas.

4 CONCLUSIONS

In this paper, a new network structure DSC Based Dual-Resunet is proposed according to the characteristics of time-frequency two-dimensional images with radio frequency interference. The network uses the residual block as feature extraction part and combines the idea of Unet architecture for two rounds of feature extraction. Compared with other existing similar methods, it speeds up convergence speed of the network, improves performance, and can identify RFI well. At the same time, the addition of depthwise separable convolution makes DSC Based Dual-Resunet have fewer parameters and FLOPs than other mainstream networks. The model size and parameters are about 12.5% of Unet, and the amount of computation is 38% of Unet but it achieves higher segmentation accuracy. This means that the network has characteristics of both lightweight and high performance. In addition, from the experimental results of SBRS, it can be seen that the recognition effect of the network is better than traditional algorithms such as SumThreshold when there are burst in the image, and it is more flexible and accurate for RFI recognition.

Although DSC Based Dual-Resunet has good performance, there are still some aspects that can be improved. In this paper, dice loss is used to solve the problem caused by imbalance of positive and negative samples in the network training process, but there is still a small chance of training failure. Therefore, the influence between loss function and characteristics of data set deserves more in-depth research in the future. At present, the network is only implemented in GPU. To better utilize the advantages of computation amount brought by depthwise separable convolution, the implementation of mobile network structure can be studied. Therefore, the optimization of the network itself and the realization of mobile network constitute the directions of future research.

References

Akeret, J., Chang, C., Lucchi, A., & Refregier, A. 2017, Astronomy and Computing, 18, 35

Bethapudi, S., & Desai, S. 2018, Astronomy and Computing, 23, 15

Chen, B., Zhao, T., Liu, J., & Lin, L. 2021, International Journal of Machine Learning and Cybernetics, 12, 1

Cendes, Y., Wijers, R. A. M. J., Swinbank, J. D., et al. 2014, arXiv e-prints, arXiv:1412.3986

Chollet, F. 2016, arXiv e-prints, arXiv:1610.02357

Czech, D., Mishra, A., & Inggs, M. 2018, Astronomy and Computing, 25, 52

Dai, C., Zuo, S. F., Liu, W., et al. 2019, in Astronomical Society of the Pacific Conference Series, 523, Astronomical Data Analysis Software and Systems XXVII, eds. P. J. Teuben, M. W. Pound, B. A. Thomas, & E. M. Warner, 71

Gong, Z., Xu, L., Tian, Z., Bao, J., & Ming, D. 2020, in 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC), 303

Guan, S., Khan, A., Sikdar, S., & Chitnis, P. V. 2018, arXiv e-prints, arXiv:1808.10848

He, K., Zhang, X., Ren, S., & Sun, J. 2015, arXiv e-prints, arXiv:1512.03385

Howard, A. G., Zhu, M., Chen, B., et al. 2017, arXiv e-prints, arXiv:1704.04861

Hu, H., Zheng, Y., Zhou, Q., et al. 2019, in 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 1197

Huang, Q., Xia, C., Wu, C., et al. 2017, arXiv e-prints, arXiv:1707.06426

Li, H., Xiong, P., An, J., & Wang, L. 2018, arXiv e-prints, arXiv:1805.10180

Liu, J., Kang, Y., Hu, D., & Chen, Y. 2020, in 2020 13th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 796

Mészárosová, H., JiřiČka, K., & Karlický, M. 1999, A&A, 348, 1005

Offringa, A. R., de Bruyn, A. G., Biehl, M., et al. 2010, MNRAS, 405, 155

Ronneberger, O., Fischer, P., & Brox, T. 2015, arXiv e-prints, arXiv:1505.04597

Srivastava, R. K., Greff, K., & Schmidhuber, J. 2015, arXiv e-prints, arXiv:1505.00387

Wolfaardt, C. J. 2016, Machine Learning Approach to Radio Frequency Interference (RFI) Classification in Radio Astronomy, PhD Thesis, Stellenbosch University

Zhang, L., Shen, J., & Zhu, B. 2020, Structural Health Monitoring, 20, 147592172094006

Zhao, J., Zou, X., & Weng, F. 2013, IEEE Transactions on Geoscience and Remote Sensing, 51, 4830