CSL-SFNet for Cooperative Spectrum Sensing in Cognitive Satellite Network with GEO and LEO Satellites

Kai Yang, Shengbo Hu, Xin Zhang, Tingting Yan, and Manqin Zhu

***Abstract*—In a cognitive satellite network (CSN) with GEO and LEO satellites, there is a large propagation losses between the sensing satellite and the ground station. The results of spectrum sensing from a single satellite may be inaccurate, which will create a serious interference in the primary satellite system. Cooperative spectrum sensing (CSS) has become the key technology to solve the above problems in recent years. However, most of the current CSS techniques are model-driven. They are difficult to model and implement in CSNs since their detection performance is strongly dependent on an assumed statistical model. Thus, we propose a novel CSS scheme, which uses convolutional neural networks (CNNs), self-attention (SA) modules, long short-term memory networks (LSTMs), and soft fusion networks, called CSL-SFNet. This scheme combines the advantages of CNNs, SA modules, and LSTMs to extract the features of the input signals from the spatial and temporal domains. Additionally, the CSL-SFNet makes use of a novel soft fusion technique that improves detection performance while also considerably reducing communication overhead. The simulation results demonstrate that the proposed algorithm can achieve a detection probability of 90% when the signal-to-noise ratio is -20 dB; it has a shorter running time and always outperforms the other CSS algorithms.**

***Index Terms*—Cognitive satellite network, cooperative spectrum sensing, deep learning.**

I. Introduction

CURRENTLY, air-space-terrestrial-sea integrated networks have become a hot research topic for B5G and 6G [1]. To achieve global coverage with low time delays, the Cubesat-based LEO mega satellite constellation, i.e., a satellite system consisting of multiple orbital planes and hundreds of small satellites, has become an important candidate technology.

However, with LEO mega satellite constellations becoming increasingly operational, the available spectrum resources are more crowded. To improve spectrum utilization, cognitive satellite networks (CSNs) with GEO and LEO satellites [2] have become important candidate technologies. In the CSN, LEO satellites are permitted to access the authorized spectrum of the GEO satellites through spectrum sensing technology, which can effectively increase the utilization efficiency of the authorized spectrum. According to [3], the GEO must be safeguarded, and the LEO's interference with the GEO must be kept to a minimum. Nevertheless, due to the large propagation delay and the fadings between the sensing satellites and the ground station [4], the spectrum sensing results of one sensing satellite may be unreliable, causing other LEO satellites to incorrectly access the authorized bands of the GEO satellites and interfere with them. Fortunately, the probability of these false sensations can be decreased by combining the sensing results from multiple space-separate LEO satellites that sense the same spectrum band, i.e., cooperative spectrum sensing (CSS) [5]. Fundamentally, CSS takes advantage of spatial separation diversity, which greatly reduces the likelihood of multiple channels deep fading at the same time [6]. Therefore, in recent years, the CSS has received an increasing amount of attention in satellite communication scenarios [7]. In CSS, the fusion center (FC) is used to make the final decision on spectrum availability. The decision-making process primarily involves either soft combining or hard combining [8]. In soft combining, each sensing satellite directly sends the local sensing data to the FC, which provides the best detection performance but at the cost of reporting link overhead [9]. The most important motivation for cognitive networks is to improve spectrum efficiency, so it is disadvantageous for cognitive networks to produce a large communication overhead in the reporting link [10]. Hard combining drastically reduces communication overhead by allowing each sensing satellite to make a local decision first, then send the decision information to the FC for the final decision. Nonetheless, this method cannot make full use of the features of the local sensing data, which will lead to a loss of features and a reduction of the detection performance [5].

Traditionally, most research on CSS has focused on model-driven methods. Their detection performance relies heavily on predefined statistical models, which makes them more difficult to model and deploy in real environments [11]. Popular detection algorithms include energy detection, matched filter detection, cyclic smooth detection, and eigenvalue detection. Among them, matched filter and cyclic smooth detection both require a priori knowledge of the primary user (PU) signals,

This work was supported by Guizhou Province Education Department Projects of China, grant number KY [2017]031, KY [2020]007, and the National Natural Science Foundation of China, grant number 61561009. (Corresponding author: Shengbo Hu, Xin Zhang)

Kai Yang is with the College of Big Data and Information Engineering, Guizhou University, Guiyang, 550025, China and with the Intelligent Information Processing Research Institute, Guizhou Normal University, Guiyang, 550001, China (email: kaiyang_21@163.com).

Shengbo Hu is with the Intelligent Information Processing Research Institute, Guizhou Normal University, Guiyang, 550001, China and with the National Space Science Center, CAS, Beijing, 100190, China (email: hsb@nssc.ac.cn).

Xin Zhang is with the College of Big Data and Information Engineering, Guizhou University, Guiyang, 550025, China (email: xzhang1@gzu.edu.cn).

Tingting Yan, and Manqin Zhu are with the Intelligent Information Processing Research Institute, Guizhou Normal University, Guiyang, 550001, China (email: 1592624854@qq.com, 1274208585@qq.com).

which is difficult to obtain in practice [12]. Although eigenvalue detection does not require corresponding a priori knowledge, it is difficult to deploy due to its high computational complexity [13]. Energy detection (ED) is simple and easy to deploy. However, its detection ability rapidly degrades or even fails to work when the signal-to-noise ratio (SNR) drops below its lower detection limit [14]. Besides, based on the energy observations of all secondary users (SUs), [15] proposed an asymptotically optimum detection algorithm. The simulation results show that the algorithm outperforms OR detector in Rayleigh fading and shadowing environments, and also it performs almost as well as the optimum detector. However, the algorithm also requires the estimations of unknown parameters in the detector structure.

Recently, with the rapid development of data-driven signal processing techniques, machine Learning (ML) and deep learning (DL) techniques have attracted extensive attention from industry and academia in the context of future wireless communication [12], [16]. Applying ML/DL to CSS can make the spectrum sensing process of each sensing satellite more adaptable to changes in the channel environment since it does not require any prior knowledge of the new environment. Additionally, the main advantage of CSN is its cognitive capability, i.e., the self-learning capability in the radio environment, which is analogous to the ML/DL model. Therefore, the ML/DL model is widely used in cognitive networks [17]. Applying ML/DL to CSS can make the spectrum sensing process of each sensing satellite more adaptable to changes in the channel environment since it does not require any prior knowledge of the new environment. [18] proposed a probabilistic spectrum sensing data falsification (SSDF) attack against a soft-judgment spectrum sensing model. In this attack, the attacker is able to perform parameter configuration adaptively. Simulation results demonstrate the effectiveness of the proposed attack strategy. For SSDF attacks scenario, [19] proposed a scheme based on clustering the secondary users to counter SSDF attacks. By using ML algorithms to classify each of the clusters as reliable or unreliable, and the simulation results showed that support vector machine (SVM) and artificial neural network (ANN) outperformed other machine learning classification algorithms and showed good detection performance. [20] proposed a method based on clustering the SUs and estimating some unknown parameters of their received power to solve the problem of SSDF attack, which is based on the clustering algorithm in classical ML, has low computational complexity and shows good detection performance in detecting malicious users. Similarly, [21] proposed a reliable method based on clustering the cooperating sensors, which significantly improves the performance of cognitive radio networks with attackers. In [22], the authors proposed a CSS model based on deep Q-learning, which outperforms the widely popular SVM-based classification methods and traditional CSS methods under SSDF attacks. [23] proposed a spectrum sensing algorithm using an SVM-optimized RBF neural network, and the results show that the detection performance of spectrum sensing can be further improved by an ML-optimized RBF neural network algorithm, which opens up a new direction for the application of ML and neural network. In [24], a federated learning (FL) algorithm is proposed to distribute the data collection and model training process over many devices. The results show that the detection accuracy of the FL algorithm is similar to that of detection using convolutional neural networks (CNNs), achieving the goal of simplifying the spectrum sensing process in the network. In [25], a machine learning-based CSS method was proposed. Although it achieved good sensing performance, it also greatly affected its robustness when the noise power was too large. In [26], ANN is used for spectrum sensing. However, ANN was prone to overfitting, which will directly affect the sensing results of the test data. In addition, [8] proposed a spatiotemporal system model of a CSS scenario for which a CNN is trained to classify the PU states to achieve spectrum sensing. Compared with traditional methods, this scheme achieves a higher detection probability. Nevertheless, the algorithm sends the local training samples directly to FC, resulting in a significant communication overhead and a decline in the efficiency of training and testing.

However, as far as we know, there are very few studies on applying ML/DL to the spectrum sensing problem in CSNs. [27] proposes a satellite-based spectrum prediction system that constructs a joint LSTM-ARMA assisted spectrum prediction scheme by combining Long Short-Term Memory networks (LSTMs) and Autoregressive Moving Average (ARMA) models. This scheme effectively reduces prediction errors and enables advanced prediction of future spectrum occupancy. In [16], a spectrum sensing method utilizing temporal convolutional networks (TCN) is introduced for CSNs. This method employs TCN to extract the time domain characteristics of the received signal in order to ascertain whether the PU exists. For a CSN with low SNR, [28] proposed a spectrum sensing scheme that combines CNNs and LSTMs, demonstrating good sensing capability. Nevertheless, these studies have only considered the spectrum sensing of individual sensing nodes and have not taken into account the fading effects on satellite-to-ground links. In this paper, we explore a novel DL-based CSS scheme in a CSN consisting of GEO and LEO satellites. In Fig. 1, the GEO relay satellite for full-climate and full-time tracking, telemetry, and command systems (TT&C) is used as the PU, and LEO satellites are used as secondary users (SUs) to share the spectrum with the PU. Between the satellite-ground link of this CSN, similar to [29], the fading model we consider mainly includes free-space transmission loss, cloud attenuation and atmospheric absorption. The main research contributions of this paper are summarized as follows:

- First, in the proposed CSN, we analyze the effects of the antenna direction of the GEO earth station on the SNR over Ka-band satellite-ground links.
- Second, in the proposed CSN, we propose a novel CSS model, which uses CNNs, self-attention (SA) modules, LSTMs, and soft fusion networks, called CSL-SFNet. This model uses a compromise between soft combining and hard combining in the FC, i.e., each SU satellite sends only local sensing soft features (SSFs) to the FC. This approach fully utilizes the feature information contained in the PU signals while simultaneously lowering the

communication overhead. The CSL-SFNet consists of CSLNet deployed on LEO satellites and Soft-FusionNet deployed on GEO satellites. The CSLNet consists of the CNN, the self-attention (SA) module, and the long short-term memory network (LSTM), which can extract the complex features of the PU signals from the spatial and temporal domains. In addition, Soft-FusionNet is mainly composed of three dense layers, which are used to aggregate SSFs from all the SUs and make the final decision. The simulation results demonstrate that the proposed algorithm can obtain a detection probability (P_d) of 90% when the SNR is -20 dB with less running time and always outperforms the other CSS algorithms.

- Finally, we must maintain a constant false alarm probability (P_f) in order to comply with the IEEE 802.22 standard's transmission requirement [30]. Therefore, to maximize P_d at a given P_f, inspired by the Neyman-Pearson criterion, this paper designs a threshold-based detection mechanism at the output of Soft-FusionNet, which can conveniently control P_f.

The remainder of this paper is organized as follows. In Section 2, the CSN system model is described. In Section 3, the proposed CSS network model framework is introduced. To evaluate the proposed algorithms, Section 4 presents the simulation results. Finally, Section 5 makes a few conclusions.

II. System Model and Problem Formulation

A. System model

In the CSN shown in Fig. 1, the GEO relay satellite working in the Ka-band is the PU, which is considered an FC because of its wide coverage and relatively static characteristics with the ground station, and the Soft-FusionNet is deployed on the PU. In addition, LEO satellites with an orbital altitude of 550 km are SUs, and CSLNet is deployed on each SU. The PU satellite beam coverage includes both multiple GEO earth stations (GESs) and LEO earth stations (LESs). The antenna beam of the SU satellite overlaps with that of the PU satellite. The PU satellite will suffer serious interference if the SU satellites access the same frequency range. Additionally, in this CSN, the PU satellite sends data to the GES in the downlink. In the uplink, the GES sends commands to the PU satellite. We assume that the spectrum data from the PU satellite or GES can be received by the SU satellites within the beam coverage of the PU satellite. For convenience, we focus on the uplink scenario where the GES communicates with a single PU satellite. During sensing, the SU satellites at different locations receive data from the GESs, and CSLNet extracts sensing soft features (SSFs) from the data first. Second, each SU satellite sends the local SSFs to the FC through the reporting link (i.e., the channel used when the SU communicates with the FC). Then, the FC combines all the SSFs and sends them to Soft-FusionNet to make the final decision. Finally, the FC broadcasts the sensing results to all the SUs. Thus, the CSS problem can be described as a binary hypothesis testing

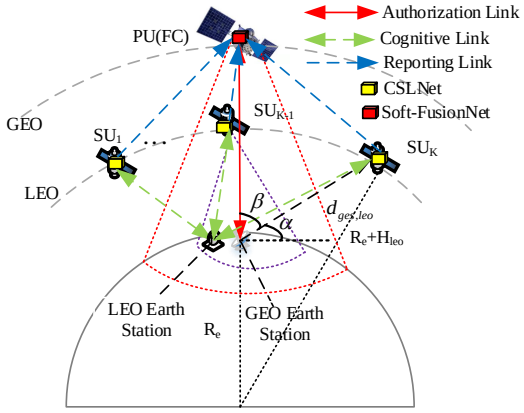

Fig. 1. Cooperative spectrum sensing model for the proposed CSN.

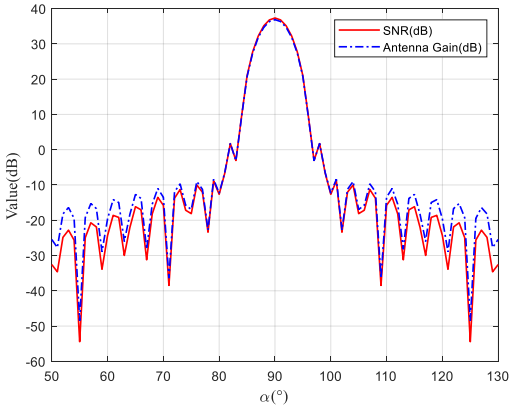

Fig. 2. Influence of elevation angle α on SNR and antenna gain.

problem, and the signal samples received by the SU satellites can be written as [29]:

$$\mathbf{x}(n) = \begin{cases} \mathbf{w}(n), & H_0 \\ \sqrt{P_{ges}^t}\sqrt{h_{ges,leo}}\mathbf{s}(n) + \mathbf{w}(n), & H_1 \end{cases}, \quad (1)$$

where

$$h_{ges,leo} = \frac{P_{leo}^r}{P_{ges}^t} = G_{ges}^t(\beta) G_{leo,max}^r \cdot \left(\frac{c}{4\pi f d_{ges,leo}(\alpha)}\right)^2 10^{-\frac{A_c + A_g}{10}}, \quad (2)$$

with H_0 and H_1 denoting the hypotheses that a PU is present and absent, respectively. $\mathbf{x}(n) \in \mathbb{C}^{N\times 1}$ denotes the signal samples received by the SU, and $\mathbf{s}(n) \in \mathbb{C}^{N\times 1}$ is the signal sample sent by the GES. $\mathbf{w}(n) \in \mathbb{C}^{N\times 1}$ represents the noise vector, which is assumed to follow a circularly symmetric complex Gaussian (CSCG) distribution with zero mean and variance σ_w^2. P_{leo}^r and P_{ges}^t denote the received power of the SU and the transmitted power of the GES, respectively. α stands for the elevation angle between the SU satellite and GES. $d_{ges,leo}(\alpha)$ is the distance of the GES to the SU, which is mainly affected by α. c and f represent the velocity of the light and the working carrier frequency, respectively. A_g and A_c represent the propagation factors of the atmospheric absorption and the cloud attenuation in the uplink, respectively

[29]. Furthermore, $G^r_{leo,max}$ is the max gain of the receive antenna of the SU satellites, which is constant because the receive antenna continuously tracks the GES. β is the off-axis angle of the GES in the direction of the SU satellites. $G^t_{ges}(\beta)$ stands for the gain of the GES transmit antenna in the direction of the SU, which is mainly affected by the β. The antenna gain is expressed by [3]:

$$G^t_{ges}(\beta) = G^t_{ges,max}\left[\frac{J_1(\mu)}{2\mu} + 36\frac{J_3(\mu)}{\mu^3}\right]^2, \tag{3}$$

where

$$\mu = 2.07123\frac{\sin(\beta)}{\sin(\beta_{3dB})}, \tag{4}$$

with $J_1(\cdot)$ and $J_3(\cdot)$ are the first- and third-order Bessel functions, respectively. β_{3dB} is the angle that corresponds to the 3 dB beamwidth. $G^t_{ges,max}$ is the maximum gain of the GES transmit antenna when $\beta = 0°$, and its expression is:

$$G^t_{ges,max} = \eta\frac{4\pi A}{(c/f)^2}, \tag{5}$$

where η stands for the antenna efficiency and A is the antenna area. In addition, it can be proven by geometric relations that β is also a function of α. The dotted line in Figure 2 presents the variation in the antenna gain $G^t_{ges}(\alpha)$ with α. The setting of simulation parameters is shown in Table II.

Then, the SNR of the received signals can be expressed as [29]

$$\begin{aligned} SNR = &\frac{P^t_{ges}G^t_{ges}(\alpha)G^r_{leo,max}}{kT_{leo}B} \\ &\cdot\left(\frac{c}{4\pi f d_{ges,leo}(\alpha)}\right)^2 10^{-\frac{A_g+A_c}{10}}, \end{aligned} \tag{6}$$

where k stands for Boltzmann's constant. B and T_{leo} denote the bandwidth and noise temperature of the receiver at the SU satellites, respectively. Fig. 1 shows that $d_{ges,leo}(\alpha)$ follows the law of first decreasing and then increasing when the sensing LEO satellite travels along its orbit at high speed. $d_{ges,leo}(\alpha)$ can be expressed as

$$d_{ges,leo}(\alpha) = \sqrt{R_e^2\sin\alpha + 2H_{leo}R_e + H_{leo}^2} - R_e\sin\alpha, \tag{7}$$

where R_e denotes the radius of the Earth and H_{leo} is the altitude of the LEO. Apparently, the SNR is mainly influenced by $G^t_{ges}(\alpha)$ and $d_{ges,leo}(\alpha)$ when f is constant. The solid line in Fig. 2 shows that the SNR of the received signals is greatly affected by the position of the SU satellites, and it can be seen that the fluctuation of SNR reaches 60 dB. Therefore, the sensing results of the SU satellites far away from the GES are likely to be unreliable. To improve the overall sensing performance, a robust CSS scheme using CSL-SFNet is designed in this paper and will be described in detail later.

III. CSL-SFNet for Cooperative Spectrum Sensing

A. CSL-SFNet

Fig. 3 shows the network architecture of the proposed CSL-SFNet for CSS, which consists of two parts. The first part, called CSLNet, is deployed on each SU satellite to obtain the SSFs from the input data. We first use a CNN to extract features from the input signals. Since the gradient of the traditional stacked CNN model will disappear as the network depth increases, we introduce the residual module, which can successfully address the problem of gradient disappearance by fusing the input with the output features of the convolution layer. Additionally, LSTM is used to extract the deeper time-dependent features from the high-dimensional abstract features output by the residual module. The main goal of the self-attentive (SA) module [31] is to focus on specific regions with strong dependencies in the output features of the residual module, which can significantly increase how efficiently crucial feature information is extracted. The second part of the architecture, called Soft-FusionNet, is deployed on a PU satellite to make the final decisions; it consists of three dense layers with 128, 32, and 2 neurons. The specific parameter settings of the proposed CSS architecture are shown in Table I.

B. Data preprocessing

For the proposed CSLNet, which is only suitable for non-complex numbers, we partition the input samples into real and imaginary parts. Thus, the input samples can be written as

$$\begin{aligned} \mathbf{x}_m = \{&[re\left(x_m(0)\right), im\left(x_m(0)\right)]; [re\left(x_m(1)\right), im\left(x_m(1)\right)]; \\ &\ldots [re\left(x_m(N)\right), im\left(x_m(N)\right)]\}, m = 1, 2, \ldots M, \end{aligned} \tag{8}$$

where $re(\cdot)$ and $im(\cdot)$ represent the real part and the imaginary part, respectively. N and M denote the length and number of samples, respectively. The received samples need to be labeled in the offline training stage, and the labeled samples can be represented as

$$\mathbf{X} = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \cdots (\mathbf{x}_M, y_M)\}\,(m = 1, 2, \cdots, M), \tag{9}$$

where $(\mathbf{x}_m, y_m)$ stands for the m-th sample of the training set $\mathbf{X}$. $y_m \in \{0, 1\}$ represents the label of $\mathbf{x}_m$. $y_m = 1$ and $y_m = 0$ denote the hypotheses of H_1 and H_0 in (1), respectively. Since CSS is a binary hypothesis test problem, training CSLNet can be considered a binary classification problem. Therefore, we can encode the label y_m as a one-hot vector

$$y_m = \begin{cases} [1, 0]^T, H_1 \\ [0, 1]^T, H_0 \end{cases} \tag{10}$$

to indicate the state of the PU.

C. Offline training

For the proposed CSS model, the offline training is divided into two stages. The first stage is that each SU trains the CSLNet through the labeled training samples offline. For convenience, we define a clear physical meaning for the network output that is normalized by the $softmax$ function, i.e.,

$$p^k_{\mathbf{\Theta}}(\mathbf{x}_m) = \begin{bmatrix} p^k_{\mathbf{\Theta}|H_1}(\mathbf{x}_m) \\ p^k_{\mathbf{\Theta}|H_0}(\mathbf{x}_m) \end{bmatrix}, \tag{11}$$

with

$$p^k_{\mathbf{\Theta}|H_1}(\mathbf{x}_m) + p^k_{\mathbf{\Theta}|H_0}(\mathbf{x}_m) = 1, \tag{12}$$

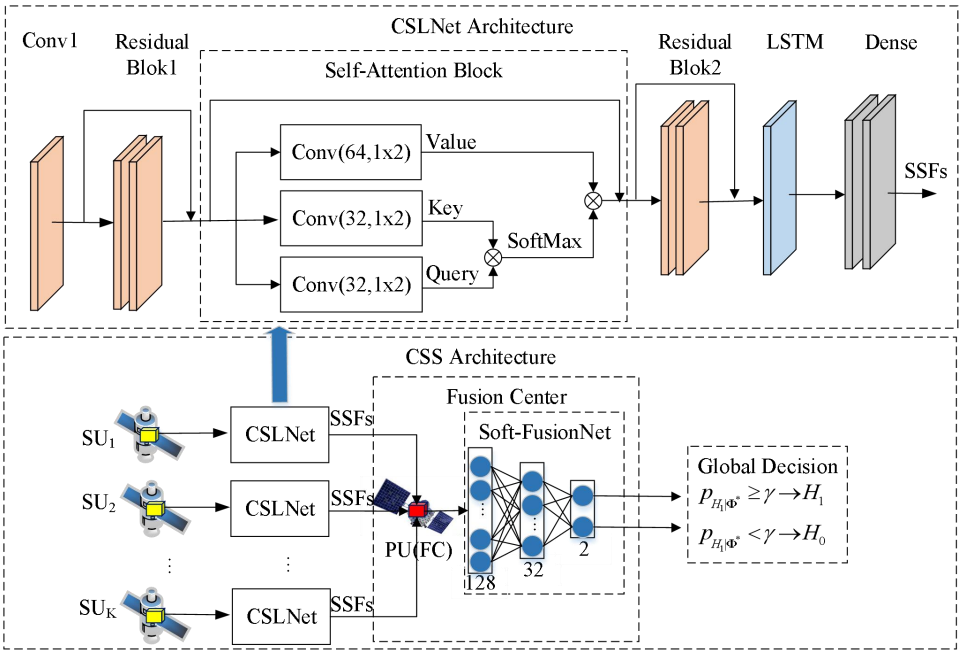

Fig. 3. The architecture of cooperative spectrum sensing in cognitive satellite network based on CSLNet. The architecture consists of two sub-network models. The front sub-network is SCLNet, which is used to extract the sensing soft features (SSFs) of the data from each SU satellite and sends the SSFs to the PU satellite for feature fusion. The rear sub-network is Soft-FusionNet, which is responsible for fusing the SSFs from various SU satellites and outputs the final spectrum sensing results.

where $\mathbf{\Theta}$ stands for the trained weights of CSLNet, and $p^k_{\mathbf{\Theta}}(\cdot)$ denotes the entire CSLNet deployed on the k-th SU satellite. $p^k_{\mathbf{\Theta}|H_i}(\mathbf{x}_m)$ represents the class score of H_i output by CSLNet. Besides, a generic objective function is defined for each SU in the training stage by using the maximum-likelihood criterion, i.e.,

$$L_k(\mathbf{\Theta}) = \prod_{m=1}^{M} \left(p^k_{\mathbf{\Theta}|H_1}(\mathbf{x}_m)\right)^{y_m} \left(p^k_{\mathbf{\Theta}|H_0}(\mathbf{x}_m)\right)^{1-y_m}. \quad (13)$$

According to objective function (13), we can define a cross-entropy cost function for CSLNet training, i.e.,

$$\begin{aligned} Loss_k(\mathbf{\Theta}) = -\frac{1}{M}\sum_{m=1}^{M} \Big[& y_m \log\left(p^k_{\mathbf{\Theta}|H_1}(\mathbf{x}_m)\right) \\ & + (1-y_m)\log\left(1-p^k_{\mathbf{\Theta}|H_1}(\mathbf{x}_m)\right)\Big]. \end{aligned} \quad (14)$$

CSLNet training's purpose is to determine the optimum weight parameter

$$\mathbf{\Theta}^* = \arg\min_{\mathbf{\Theta}} Loss_k(\mathbf{\Theta}). \quad (15)$$

We obtain the optimal model parameter $\mathbf{\Theta}^*$ by minimizing (14). Meanwhile, the backpropagation algorithm and the Adam-based optimizer are used to train the CSLNet. Then, we call the output of the well-trained CSLNet the SSF and express it as

$$p^k_{\mathbf{\Theta}^*}(\mathbf{x}_m) = \begin{bmatrix} p^k_{\mathbf{\Theta}^*|H_1}(\mathbf{x}_m) \\ p^k_{\mathbf{\Theta}^*|H_0}(\mathbf{x}_m) \end{bmatrix}. \quad (16)$$

In the second stage, FC first combines all the SSFs sent by the SU satellites into one SSF vector

$$\mathbf{P}_{SSF} = \left\{p^1_{\mathbf{\Theta}^*}(\mathbf{x}_m); p^2_{\mathbf{\Theta}^*}(\mathbf{x}_m); \cdots p^K_{\mathbf{\Theta}^*}(\mathbf{x}_m)\right\}. \quad (17)$$

Second, $\mathbf{P}_{SSF}$ is used to train Soft-FusionNet to obtain the optimal model parameter $\mathbf{\Phi}^*$. Since the training process of Soft-FusionNet is similar to that of CSLNet, we will not describe it in detail. Finally, the outputs of the well-trained Soft-FusionNet are $p_{\mathbf{\Phi}^*|H_1}(\mathbf{x}_m)$ and $p_{\mathbf{\Phi}^*|H_0}(\mathbf{x}_m)$, which are the class probabilities of H_1 and H_0, respectively.

Finally, Fig. 4 presents the loss value changes during the offline training process of the single CSLNet and Soft-FusionNet. We can clearly see that the loss values of both networks decrease as the number of epochs increases, without any apparent overfitting. Moreover, we can also observe that the loss value of Soft-FusionNet is significantly lower than that of CSLNet. This indicates that the SSFs extracted by multiple CSLNets can effectively reduce the decision error in FC through soft fusion. This result demonstrates the effectiveness of the proposed soft fusion scheme in CSS.

D. Online detection

In the online detection part, the SU satellites receive new unlabeled signal samples from GES and send them into the proposed CSS architecture to obtain the corresponding class probabilities $p_{\mathbf{\Phi}^*|H_1}(\mathbf{x}_m)$ and $p_{\mathbf{\Phi}^*|H_0}(\mathbf{x}_m)$. We can obtain the final decision result by comparing the values of $p_{\mathbf{\Phi}^*|H_1}(\mathbf{x}_m)$ and $p_{\mathbf{\Phi}^*|H_0}(\mathbf{x}_m)$. However, to fulfill the IEEE 802.22 standard's transmission requirement, we must maintain a constant P_f. Inspired by the Neyman-Pearson criterion, we design a threshold-based decision scheme at the output of Soft-FusionNet that can control the desired P_f by updating the threshold. First, we randomly choose L samples under the $y_m = 0$ label from the training samples and constitute the selected L samples as a new sample set $\{\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \cdots \tilde{\mathbf{x}}_L\}$.

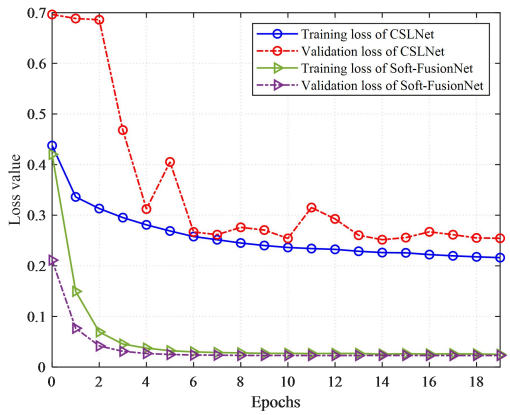

Fig. 4. Training loss of CSLNet and Soft-FusionNet.

Algorithm 1 CSL-SFNet Based CSS.

1: Set $i = 0$; initialize maximum training epochs E; initialize $\boldsymbol{\Theta}$ and $\boldsymbol{\Phi}$ with random weights;
2: Collect the training set $\mathbf{X}$ and feed it into CSLNet;
3: **for** $i \leq E$ **do**
4: According to (15), SUs train the CSLNet by the backpropagation algorithm to obtain the optimal model parameters $\boldsymbol{\Theta}^*$;
5: **end for**
6: SUs send the output SSFs of well-trained CSLNet to FC;
7: FC combines SSFs as $\mathbf{P_{SSF}}$;
8: **for** $i \leq E$ **do**
9: FC trains the Soft-FusionNet by the backpropagation algorithm to get the optimal model parameters $\boldsymbol{\Phi}^*$;
10: **end for**
11: FC calculates the threshold γ based on (18) and (19);
12: SUs receive the test sample $\mathbf{x}'_m$ online and input it into the well-trained CSLNet to obtain the new SSFs, which are transmitted to FC;
13: FC outputs the corresponding category probability vector $p_{\boldsymbol{\Phi}^*|H_1}(\mathbf{x}'_m)$ and $p_{\boldsymbol{\Phi}^*|H_0}(\mathbf{x}'_m)$ and decides the final PU state based on (20);
14: FC broadcasts the final decision result to each SU.

Then, the sample set is put into the proposed CSS scheme to obtain the corresponding results. Then, we sort the results in descending order

$$p_{\boldsymbol{\Phi}^*|H_1}(\tilde{\mathbf{x}}_k) \geq p_{\boldsymbol{\Phi}^*|H_1}(\tilde{\mathbf{x}}_l), \quad \forall 1 \leq k \leq l \leq L. \tag{18}$$

Second, the detection threshold can be expressed as

$$\gamma = p_{\boldsymbol{\Phi}^*|H_1}\left(\tilde{\mathbf{x}}_{round(P_f L)}\right), \tag{19}$$

where $rund(\cdot)$ denotes the rounding down function. Finally, for the new online samples $\mathbf{x}'_m$, the decision result can be obtained by

$$p_{\boldsymbol{\Phi}^*|H_1}(\mathbf{x}'_m) \underset{H_0}{\overset{H_1}{\gtrless}} \gamma. \tag{20}$$

Above, the proposed CSS algorithm is summarized in algorithm 1.

IV. Simulation and Analysis

A. Simulation environments

In the proposed CSN scenario, we use the GNURadio [32] to generate the sample dataset of PU signals. The PU signal is a 16QAM modulated signal commonly used in satellite communication scenarios [14]. The SNR of the received signals is greatly influenced by α. Thus, to verify the SNR robustness of the proposed algorithm, we generate weak signal samples in the SNR range of -20 dB $\sim$ 0 dB as the dataset according to (6). The relevant parameter settings for the generated data samples are shown in Table II. In addition, to further analyze the effects of the direction of the antenna of GES on the SNR over the Ka-band satellite-ground links, we incorporate noise uncertainty (NU), which has a significant impact on the performance of the actual detection. For the NU scenario, the estimated noise power is $\hat{\sigma}_w^2 = \varepsilon\sigma_w^2$, where σ_w^2 denotes the actual noise power. ε stands for the NU factor [11], which obeys a uniform distribution with an interval of $[-U, U]$. In our scenario, we let $U = 1$, which indicates that 1 dB of NU is introduced. Finally, all the algorithms are implemented based on Python 3.7, and the simulation platform is a PC equipped with an NVIDIA GeForce RTX2080Ti GPU and an Intel(R) Core i9-9900K CPU.

B. Simulation results

In this part, the performance of all the CSS algorithms will be evaluated by two key performance indicators, i.e., P_f and P_d, which can be represented as $P_f = P(T > \gamma^* \mid H_0)$ and $P_d = P(T > \gamma^* \mid H_1)$, respectively, where T and γ^* are the test statistics and detection threshold of the corresponding algorithm, respectively.

First, we will verify the impact of the number of SU satellites on the detection performance of the proposed algorithm. We set different SU numbers and selected test samples with SNR = -16 dB to evaluate the detection performance of the proposed scheme. In Fig. 5, we clearly discover that when the SU number increases, the detection performance of the proposed approach also enhances. This is because more hidden features of the PU signals can be learned by the proposed algorithm as the number of SUs increases. In addition, considering the operating point with $P_f = 1\%$, the P_d of the proposed algorithm reaches 99% when the number of SUs is 7, but only 25% in the case of a single SU. Meanwhile, we also find that the P_d of the proposed scheme is close to convergence when the number of SUs is 7.

Second, to further evaluate the detection performance of the proposed algorithm, we compare it with other spectrum sensing algorithms, including ED, CNN, LSTM, and multilayer perceptron (MLP). The CNN model consists of two convolution layers containing 128, 64 filters with a filter size of 3 and two dense layers with 128, 2 neurons. The LSTM model consists of one LSTM layer containing 64 neurons and two dense layers with 64, 2 neurons. And the MLP consists of five dense layers with 256, 500, 250, 120, and 2 neurons, respectively. All the algorithms are evaluated using data samples with an SU number of 9 and a sample length of 512 to ensure a fair comparison, and the optimal

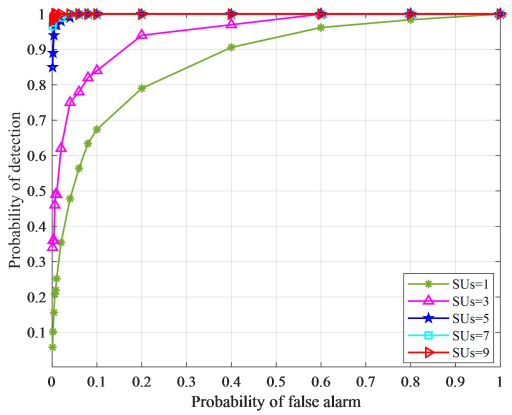

Fig. 5. ROC curves for different numbers of SU satellites with SNR=-16 dB.

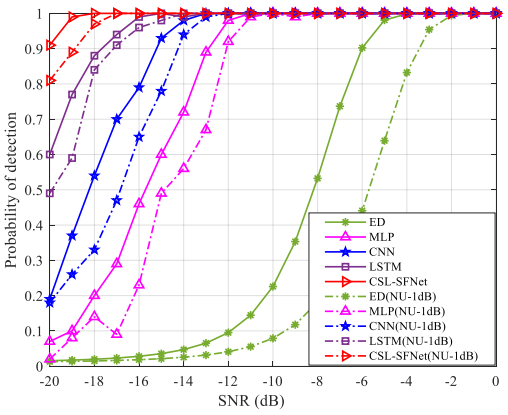

Fig. 7. P_d-SNR curves of different algorithms with SUs = 9 and P_f = 1%.

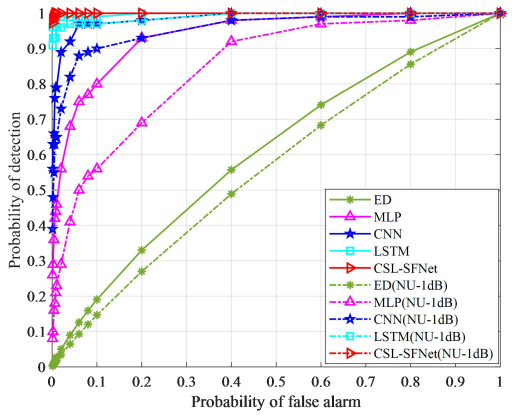

Fig. 6. ROC curves of different algorithms with SUs = 9 and SNR = -16 dB.

hyperparameters of each model are determined using a large number of simulations. According to (20), different detection thresholds can be obtained by changing the value of P_f. We choose the test sample data with SNR = -16 dB under various thresholds and feed it into each detection algorithm to obtain the receiver operating characteristics (ROC) curves. In Fig. 6, we find that the proposed algorithm outperforms other schemes at any P_f level. Apparently, the data-driven deep learning algorithms are better than the model-driven ED. This is because the neural network model has adaptive learning ability and can automatically extract hidden features from the PU signals, but ED can only make decisions based on the energy level of the PU signals. When the SNR level is low, it is difficult for the ED to accurately sense the state of the PU. In addition, NU is also common in CSNs, so it is necessary to study the robustness of the SNR of the proposed algorithm in the case of NU. The dotted line in Fig. 6 presents the sensing performance of each detection algorithm when the NU is 1 dB. We find that the proposed algorithm is less affected by NU, while the detection performance of other schemes is more affected by NU. Additionally, to verify the effectiveness of the proposed scheme at different SNR levels, we set P_f to 1% and show the P_d- SNR curves of each algorithm in Fig. 7. As with the ROC curve, the proposed algorithm still outperforms other sensing algorithms, achieving a P_d of over 90% even at SNR = -20 dB, while the ED can hardly work at this SNR level. Then, to further verify the sensing capability of the proposed algorithm in different SNR ranges, we generated untrained test samples with a SNR range of -28 to -20 dB for testing all sensing algorithms. As can be seen in Fig. 8, the proposed algorithm still shows the best detection performance at these untrained SNR levels, while CNN, MLP, and ED can barely work.

Additionally, to demonstrate the effectiveness of each component in the proposed CLS-SFNet, we further conducted ablation experiments. From the Fig. 9, we can see that the impact of the LSTM layer on sensing performance is relatively significant. This is because the model's input IQ data is a type of time series, and the LSTM can further learn the temporal correlations within the input features, thereby significantly improving the perception performance. However, the impact of the SA module on sensing performance is relatively smaller, but SA can help the model weight the data at different positions in the feature sequence when processing it, allowing the model to focus on the most important parts of the sequence, thereby improving the efficiency of model training. Overall, both the LSTM layer and the SA are essential components of the proposed model, and are crucial for the enhancement of the overall model performance.

Finally, we further evaluate the generalization ability of the proposed approach using the RadioML2016.10a [32] dataset, which contains 8 types of modulation signals and is widely used for modulation recognition scenarios. It consists of the I/Q signal vector of size $2 \times 128 \times 1$, which takes into account typical radio communication impairments such as time delays, frequency offsets, and sample rate drifts. In Fig. 10, the dotted line represents the detection ability of the proposed algorithm for untrained modulation schemes. We find that the P_d of the proposed approach for all the modulated signals reaches 100% when SNR = -12 dB. The results show that the proposed algorithm still has good detection abilities for untrained modulated signals.

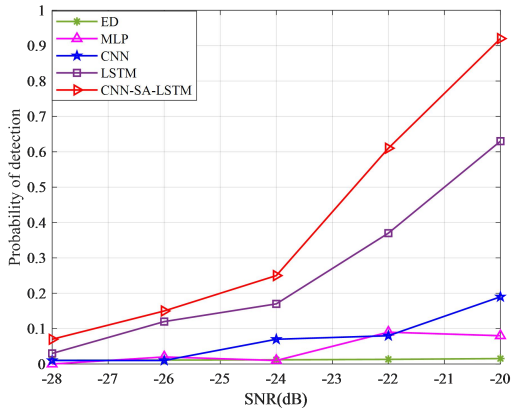

Fig. 8. Sensing performance at different SNR ranges with SUs = 9 and P_f = 1%.

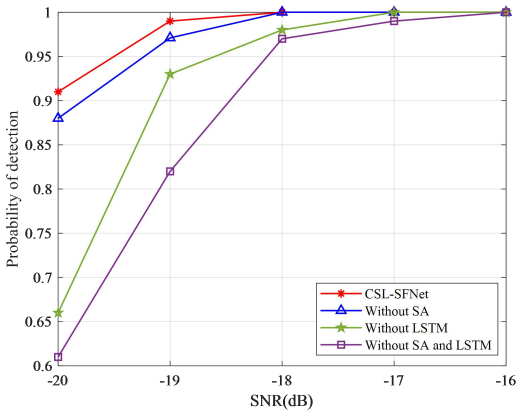

Fig. 9. Ablation experiments with SUs = 9 and P_f = 1%.

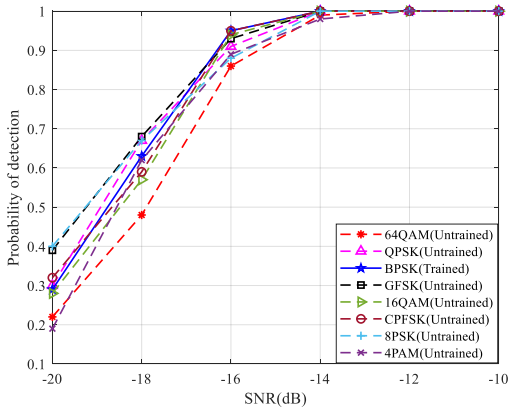

Fig. 10. P_d-SNR curves of different modulation types with SUs = 9 and P_f = 1%.

C. Complexity analysis

The proposed CSL-SFNet consists of a CNN layer, LSTM layer, and dense layer. For the CNN layer, the computational complexity of one data sample is $O\left(\sum_{i=1}^{I} n_{c,i-1} s_{c,i} n_{c,i} m_{c,i}\right)$ [33], where I is the number of convolution layers; $n_{c,i}$ is the number of convolution kernels in the i-th layer; $n_{c,i-1}$ is the number of input channels in the i-th layer; $s_{c,i}$ is the spatial size of convolution kernels; and $m_{c,i}$ is the spatial size of the output features of the i-th layer. According to Table I, the number of convolution kernels of the CNN layers used is only 64 and 32. For simplicity, we let $n_{c,1} = 64$ and $n_{c,7} = 32$ represent the other CNN layers with the same number of convolution kernels. In addition, $s_{c,1} = 2$ is used to represent the convolution kernel size of all the other convolution layers. In our spectrum sensing, the size of one sample is $2 \times N \times 1$, so the total computational complexity of the convolution part is $O\left(s_{c,1} n_{c,1} N \left(2 + \frac{5 n_{c,1}}{2} + n_{c,7}\right)\right)$. According to [34], for the LSTM layer, its computational complexity is related to the number of hidden units and the input size. In our scenario, the computational complexity of the LSTM layer is $O\left(4 n_l \left(\frac{N}{2} + n_l + 1\right)\right)$, where n_l stands for the number of hidden units. Additionally, for the dense layer, the computational complexity is $O\left(n_{in} n_{d,i}\right)$, where n_{in} and $n_{d,i}$ represent the input size and the number of neurons in the dense layer, respectively. Finally, the total computational

TABLE I
CONFIGURATION OF THE CSL-SFNET MODEL

Hyper-parameter	**Layer name**
CNN (Activation function: $ReLu$**)**	
64×(1×2)	Conv1
(1×2)	Maxpool1
64×(1×2)	Residual Blok1: Conv2,3
64×(1×2)	Residual Blok2: Conv4,5
64×(1×2)	Value: Conv6
32×(1×2)	Key & Query: Conv7,8
LSTM (Activation function: $ReLu$**)**	
Number of hidden units: 64	LSTM1
Dense (Activation function: $Softmax$**)**	
Neurons number: 32,2	Dense1,2
Neurons number: 128,32,2	Soft-FusionNet:Dense3,4,5

TABLE II
DESCRIPTION OF THE SYSTEM MODEL PARAMETERS

Parameter	**Symbol**	**Value**
Training set size (per SNR)	-	1000
Test set size (per SNR)	-	100
Batch size	-	512
Dropout rate	-	0.1
Learning rate	-	0.0001
Optimizer	-	Adam
Modulation scheme	-	16QAM
SNR range	-	-20∼0dB
Sample length	N	512
Transmit power of GES	P_{ges}^{t}	40dBm
Noise temperature of SU	T_{leo}	175K
Carrier frequency	f	29.9GHz
Bandwidth	B	24MHz
Atmospheric absorption factor	A_g	0.75dB
Cloud attenuation factor	A_c	1.25dB
Antenna diameter of SU	D_{leo}	0.3m
Antenna diameter of GES	D_{ges}	0.3m
Antenna efficiency	η	55%
Probability of detection/false alarm	P_d/P_f	-
Optimal model parameter of CSLNet	$\mathbf{\Theta}^*$	-
Optimal model parameter of Soft-FusionNet	$\mathbf{\Phi}^*$	-
Maximum gain of the GES's transmit antenna	$G_{ges,max}^{t}$	-
Maximum gain of the SU's receive antenna	$G_{leo,max}^{r}$	-
Noise uncertaint factor	ε	-

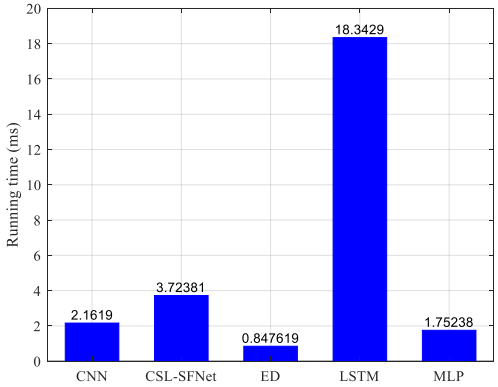

Fig. 11. Running time of different algorithms.

complexity of the proposed CSS architecture is

$$O\Big(s_{c,1}n_{c,1}N\left(2+\frac{5n_{c,1}}{2}+n_{c,7}\right)+4n_l\left(\frac{N}{2}+n_l+1\right)+ n_l n_{d,1}+n_{d,1}n_{d,2}+n_{d,2}n_{d,3}+n_{d,3}n_{d,4}+n_{d,4}n_{d,5}\Big).$$

Outside of the computational complexity, the communication overhead between the SU satellites and the FC is also an important problem. In the proposed scheme, the SU satellite sends the SSFs to the FC through the reporting link instead of directly sending the received signal data to the FC. The communication overhead required by the SU to send the SSFs of one data sample is only 96 bytes, while the communication overhead required to send one data sample to the FC is 8000 bytes. In the simulation, the total number of training samples is 42,000 and the number of SUs is 9. The proposed algorithm converges after training for 20 epochs, and the total communication overhead is only 692.14 Mbytes. However, the communication overhead required is 59062.5 Mbytes if all the SUs send data samples directly to the FC. Evidently, the proposed algorithm will show impressive advantages in future scenarios with larger data.

Finally, we analyze the test samples using different algorithms and calculate the running time accordingly. Analyzing Fig. 6 and Fig. 11, we find that when SNR = -20 dB, the CSS algorithm based on LSTM and the proposed algorithm can achieve more than 90% of P_d. However, the running time of the algorithm based on LSTM is almost five times that of the proposed algorithm. Because LSTMs have a more complex internal structure, leading to slower processing speeds for longer input data. Although proposed model includes LSTM layers, their inputs are feature data that has been dimensionality reduced by multiple convolutional modules, resulting in relatively lower complexity. This enables the proposed model to have shorter running times. In addition, the longer running time is very disadvantageous for the spectrum sensing of the high-speed moving LEO satellites. Thus, although the running time of the proposed approach is slightly longer than that of a traditional ED, the proposed scheme is a promising CSS scheme in terms of its excellent detection performance.

V. Conclusion

In this paper, we propose a novel CSS algorithm using CSL-SFNet in a CSN with LEO and GEO satellites. The algorithm employs a soft fusion technique in FC, which can fully utilize the features of the primary signals and significantly lower the communication overhead of the reporting link. The simulation results show that the proposed algorithm is insensitive to noise uncertainty and is robust to untrained modulation types. Furthermore, even in an environment with SNR = -20 dB, the P_d of the proposed algorithm can still reach more than 90%, and it always outperforms the other CSS algorithms with less communication overhead and a shorter running time.

VI. Author contributions

Shengbo Hu and Kai Yang: Methodology. **Xin Zhang and Kai Yang:** Software. **Tingting Yan and Manqin Zhu:** Validation. **Kai Yang and Tingting Yan:** Formal analysis. **Xin Zhang and Kai Yang:** Investigation. **Shengbo Hu and Kai Yang:** Data curation. **Xin Zhang:** Visualization. **Shengbo Hu and Xin Zhang:** Supervision. **Shengbo Hu:** Project administration and funding acquisition.

VII. Acknowledgements

This research was supported by Guizhou Province Education Department Projects of China, grant number KY [2017]031, KY [2020]007, and the National Natural Science Foundation of China, grant number 61561009.

VIII. Conflict of interest

The author declares that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

IX. Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

X. Orcid

Shengbo Hu: https://orcid.org/0000-0002-7891-2451

References

[1] X. You, C.-X. Wang, J. Huang, X. Gao, Z. Zhang, M. Wang, Y. Huang, C. Zhang, Y. Jiang, J. Wang, M. Zhu, B. Sheng, D. Wang, Z. Pan, P. Zhu, Y. Yang, Z. Liu, P. Zhang, X. Tao, S. Li, Z. Chen, X. Ma, C.-L. I, S. Han, and Y.-C. Liang, "Towards 6g wireless communication networks: Vision, enabling technologies, and new paradigm shifts," *Science China Information Sciences*, vol. 64, pp. 1–74, 2021.

[2] R. Liu, S. Zhu, and C. Li, "Review of cognitive satellite communication technology," in *2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, vol. 9, pp. 1378–1385, IEEE, 2020.

[3] C. Wang, D. Bian, S. Shi, J. Xu, and G. Zhang, "A novel cognitive satellite network with geo and leo broadband systems in the downlink case," *IEEE Access*, vol. 6, pp. 25987–26000, 2018.

[4] K. Shi, X. Zhang, S. Zhang, and H. Li, "Time-expanded graph based energy-efficient delay-bounded multicast over satellite networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 9, pp. 10380–10384, 2020.

[5] I. F. Akyildiz, B. F. Lo, and R. Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: A survey," *Physical communication*, vol. 4, no. 1, pp. 40–62, 2011.

[6] K. Wu, M. Tang, C. Tellambura, and D. Ma, "Cooperative spectrum sensing as image segmentation: A new data fusion scheme," *IEEE Communications Magazine*, vol. 56, no. 4, pp. 142–148, 2018.

[7] M. Jia, X. Liu, Z. Yin, Q. Guo, and X. Gu, "Joint cooperative spectrum sensing and spectrum opportunity for satellite cluster communication networks," *Ad Hoc Networks*, vol. 58, pp. 231–238, 2017.

[8] P. Shachi, K. Sudhindra, and M. Suma, "Convolutional neural network for cooperative spectrum sensing with spatio-temporal dataset," in *2020 International conference on artificial intelligence and signal processing (AISP)*, pp. 1–5, IEEE, 2020.

[9] P. M. Pradhan and G. Panda, "Information combining schemes for cooperative spectrum sensing: A survey and comparative performance analysis," *Wireless Personal Communications*, vol. 94, pp. 685–711, 2017.

[10] A. Jamshidi, "Performance analysis of low average reporting bits cognitive radio schemes in bandwidth constraint control channels," *IET communications*, vol. 3, no. 9, pp. 1544–1556, 2009.

[11] C. Liu, J. Wang, X. Liu, and Y.-C. Liang, "Deep cm-cnn for spectrum sensing in cognitive radio," vol. 37, pp. 2306–2321, IEEE, 2019.

[12] B. Soni, D. K. Patel, and M. López-Benítez, "Long short-term memory based spectrum sensing scheme for cognitive radio using primary activity statistics," *IEEE Access*, vol. 8, pp. 97437–97451, 2020.

[13] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio: State-of-the-art and recent advances," *IEEE signal processing magazine*, vol. 29, no. 3, pp. 101–116, 2012.

[14] T. Ni, X. Ding, Y. Wang, J. Shen, L. Jiang, and G. Zhang, "Spectrum sensing via temporal convolutional network," *China Communications*, vol. 18, no. 9, pp. 37–47, 2021.

[15] A. Taherpour, Y. Norouzi, M. Nasiri-Kenari, A. Jamshidi, and Z. Zeinalpour-Yazdi, "Asymptotically optimum detection of primary user in cognitive radio networks," *IET communications*, vol. 1, no. 6, pp. 1138–1145, 2007.

[16] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98–105, 2016.

[17] C. Clancy, J. Hecker, E. Stuntebeck, and T. O'Shea, "Applications of machine learning to cognitive radio networks," *IEEE Wireless Communications*, vol. 14, no. 4, pp. 47–52, 2007.

[18] A. Ahmadfard, A. Jamshidi, and A. Keshavarz-Haddad, "Probabilistic spectrum sensing data falsification attack in cognitive radio networks," *Signal Processing*, vol. 137, pp. 1–9, 2017.

[19] N. Parhizgar, A. Jamshidi, and P. Setoodeh, "Defense against spectrum sensing data falsification attack in cognitive radio networks using machine learning," in *2022 30th International Conference on Electrical Engineering (ICEE)*, pp. 974–979, IEEE, 2022.

[20] M. Ghaznavi and A. Jamshidi, "A low complexity cluster based data fusion to defense against ssdf attack in cognitive radio networks," *Computer Communications*, vol. 138, pp. 106–114, 2019.

[21] M. Ghaznavi and A. Jamshidi, "A reliable spectrum sensing method in the presence of malicious sensors in distributed cognitive radio network," *IEEE Sensors Journal*, vol. 15, no. 3, pp. 1810–1816, 2014.

[22] A. Paul, A. K. Mishra, S. Shreevastava, and A. K. Tiwari, "Deep reinforcement learning based reliable spectrum sensing under ssdf attacks in cognitive radio networks," *Journal of Network and Computer Applications*, vol. 205, p. 103454, 2022.

[23] S. Yang and C. Tong, "Cognitive spectrum sensing algorithm based on an rbf neural network and machine learning," *Neural Computing and Applications*, vol. 35, no. 36, pp. 25045–25055, 2023.

[24] M. Wasilewska, H. Bogucka, and A. Kliks, "Federated learning for 5g radio spectrum sensing," *Sensors*, vol. 22, no. 1, 2022.

[25] K. M. Thilina, K. W. Choi, N. Saquib, and E. Hossain, "Machine learning techniques for cooperative spectrum sensing in cognitive radio networks," *IEEE Journal on selected areas in communications*, vol. 31, no. 11, pp. 2209–2221, 2013.

[26] Y.-J. Tang, Q.-Y. Zhang, and W. Lin, "Artificial neural network based spectrum sensing method for cognitive radio," in *2010 6th international conference on wireless communications networking and mobile computing (WiCOM)*, pp. 1–4, IEEE, 2010.

[27] X. Ding, Q. Lv, Y. Zou, and G. Zhang, "Spectrum prediction for satellite based spectrum-sensing systems using deep learning," in *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, pp. 3472–3477, 2022.

[28] X. Ding, T. Ni, Y. Zou, and G. Zhang, "Deep learning for satellites based spectrum sensing systems: A low computational complexity perspective," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 1, pp. 1366–1371, 2022.

[29] C. Zhang, C. Jiang, J. Jin, S. Wu, L. Kuang, and S. Guo, "Spectrum sensing and recognition in satellite systems," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 2502–2516, 2019.

[30] Z. Su, K. C. Teh, S. G. Razul, and A. C. Kot, "Deep non-cooperative spectrum sensing over rayleigh fading channel," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 4, pp. 4460–4464, 2021.

[31] H. Zhao, J. Jia, and V. Koltun, "Exploring self-attention for image recognition," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10076–10085, 2020.

[32] T. J. O'shea and N. West, "Radio machine learning dataset generation with gnu radio," in *Proceedings of the GNU radio conference*, vol. 1, 2016.

[33] K. He and J. Sun, "Convolutional neural networks at constrained time cost," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 5353–5360, 2015.

[34] H. Sak, A. Senior, and F. Beaufays, "Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition," *Computer Science*, pp. 338–342, 2014.