Article

Spectrum Sensing Method Based on Information Geometry and Deep Neural Network

Kaixuan Du [1], Pin Wan [1,2], Yonghua Wang [1,3,*] , Xiongzhi Ai [1] and Huang Chen [1]

[1]   School of Automation, Guangdong University of Technology, Guangzhou 510006, China;
     2111704051@mail2.gdut.edu.cn (K.D.); wanpin2@163.com (P.W.); 17607011056@163.com (X.A.);
     gdutchenhuang@163.com (H.C.)

[2]   Hubei Key Laboratory of Intelligent Wireless Communications, South-Central University for Nationalities,
     Wuhan 430074, China

[3]   State Key Laboratory of Management and Control for Complex Systems, Institute of Automation,
     Chinese Academy of Sciences, Beijing 100190, China

*   Correspondence: wangyonghua@gdut.edu.cn

Abstract: Due to the scarcity of radio spectrum resources and the growing demand, the use of spectrum sensing technology to improve the utilization of spectrum resources has become a hot research topic. In order to improve the utilization of spectrum resources, this paper proposes a spectrum sensing method that combines information geometry and deep learning. Firstly, the covariance matrix of the sensing signal is projected onto the statistical manifold. Each sensing signal can be regarded as a point on the manifold. Then, the geodesic distance between the signals is perceived as its statistical characteristics. Finally, deep neural network is used to classify the dataset composed of the geodesic distance. Simulation experiments show that the proposed spectrum sensing method based on deep neural network and information geometry has better performance in terms of sensing precision.

Keywords: spectrum sensing; information geometry; statistical manifold; geodesic distance; deep neural network

1. Introduction

Cognitive radio (CR) is a method to alleviate the allocation of spectrum resources in response to the shortage of spectrum resources [1]. Spectrum sensing is an important technology of cognitive radio. Classical spectrum sensing algorithms include energy detection, matched filter detection, and cyclostationary characteristic detection [2]. On the basis of classical spectrum sensing, a study proposed a goodness-of-fit test algorithm [3]. The goodness-of-fit test is to study the overall distribution of the sensing signal, so this algorithm shows better performance than the classical algorithm in terms of sensing accuracy. Traditional sensing methods and goodness-of-fit tests are single secondary user (SU) methods, which are easily interfered by propagation environment factors. Therefore, some scholars applied the cooperative spectrum sensing algorithm to avoid such interference [4–6]. A sensing method based on random matrix theory (RMT) is one of cooperative spectrum sensing methods. Based on the RMT theory, a study proposed a spectrum sensing algorithm based on the maximum eigenvalue detection (MED), but this algorithm only uses the single threshold detection of the largest eigenvalue to cause the detection performance [7]. The work in [8] proposed a maximum eigenvalue-energy detection (ME-ED) spectrum sensing algorithm. The algorithm uses the ratio of the maximum eigenvalue and the average energy of the sampled signal as a statistical feature. It compares the statistical feature with the threshold to determine whether the primary user (PU) signal exists. However, the ME-ED does not participate in the calculation of the minimum eigenvalue of the sensing signal, so the information of the

sensing signal is not fully utilized. A study proposed the maximum and minimum eigenvalue (MME) spectrum sensing algorithm, which made full use of the sampling information [9]. However, if the sampling information is small, there is a large error in the acquisition of the eigenvalue, which affects the performance of the algorithm. The research work [10] proposed a spectrum sensing algorithm based on the difference between the maximum and minimum eigenvalues (DMM). The algorithm eliminates the noise component by the difference between the maximum and minimum eigenvalues, so that the same environment can show better performance.

The classical spectrum sensing methods often need to set the corresponding decision threshold according to the statistical characteristics of the eigenvalues. Next, we judge that the PU exists by the judgment criterion. When machine learning is used for classification, it has the advantages of high classification accuracy and strong learning ability. Therefore, more and more literature works have combined machine learning and spectrum sensing for classification. The work in [11] used the support vector machine method in machine learning to classify the training data. It did not need a decision threshold, but it improved the sensing accuracy. Some scholars [12] applied a combination of the k-means clustering algorithm and spectrum sensing in unsupervised learning. It achieved blind detection of the PU signal with low signal-to-noise ratio (SNR) and exhibited good sensing performance.

In this paper, the characteristics of the sensing signal matrix are extracted by way of information geometry, and the performance of spectrum sensing is improved from a new perspective. Information geometry is a set of theoretical systems developed from the study of probability distribution manifolds [13]. The concept of information geometry can be used to transform the detection problem on the spectrum into a geometric problem on the statistical manifold. The geometric method on the statistical manifold can be used to analyze the signal detection problem more intuitively [14]. In order to improve the sensing accuracy, a feature extraction method was introduced [15]. It used the geodesic distance between signals as a feature to detect signals, thus achieving the purpose of combining information geometry with spectrum sensing. The work in [16] was mainly to solve the influence of noise on the sensing precision in complex environments, and it proposed a sensing algorithm based on empirical mode decomposition (EMD) and information geometry. It used EMD to perform noise reduction processing and then used information geometry to perform feature extraction. A recent research work applied unsupervised learning to the classification of sensing signals [17]. This study combined information geometry and fuzzy C-means clustering algorithms to improve classification.

Deep learning has good migration learning ability, and its application improves the sensing performance in spectrum sensing. In [18], a blind spectrum sensing method based on deep learning was proposed. By comparing this method with an energy detector, good sensing performance under the low SNR was achieved. Xie et al. [19] proposed a spectrum sensing algorithm based on a deep learning convolutional neural network, which avoided the influence of the accuracy of the hypothetical statistical model on the detection results and improved the detection probability. Ke et al. [20] used a convolutional neural network (CNN) and the long short term memory (LSTM) methods for passive signal detection, improving signal detection performance in low SNR environments. The work in [21] proposed a hierarchical large scale neural network model, which was a hierarchical extension based on information geometry, thus improving the ability to process data.

In this paper, deep learning is applied to spectrum sensing, and the spectrum sensing method based on information geometry and deep neural network (IG-DNN) is proposed. Compared with the literature on spectrum sensing, which combines information geometry with unsupervised learning clustering algorithms, the innovation in this paper is a combination of deep learning algorithms and information geometry. The neural network algorithm in supervised learning is used to train the distance characteristics between the sampled signals to obtain a classifier. The method not only jumps out of the traditional spectrum sensing method, but also combines with deep learning to classify and thus improve the sensing effect.

The remainder of this paper is described as follows. Section 2 introduces the application scenario of spectrum sensing. Section 3 shows the method of processing spectrum sensing with information geometry. Section 4 mainly introduces deep neural networks. Section 5 is the experimental simulation. Section 6 is a summary of the paper.

2. Spectrum Sensing Model and Centralized Spectrum Sensing Scene

The SUs sense noise in a noisy environment and calculate a covariance matrix of the noise matrix. Then, the information geometry method is used to map the covariance matrix to the manifold, and the Riemann mean of the noise on the manifold is calculated. Similarly, the signal covariance matrix extracted from the sensing environment is mapped onto the manifold. The geodesic distance of the sensing signal and Riemann mean is calculated, that is the statistical characteristics of the signal are also the input data of the network. The whole process can be seen in Figure 1.
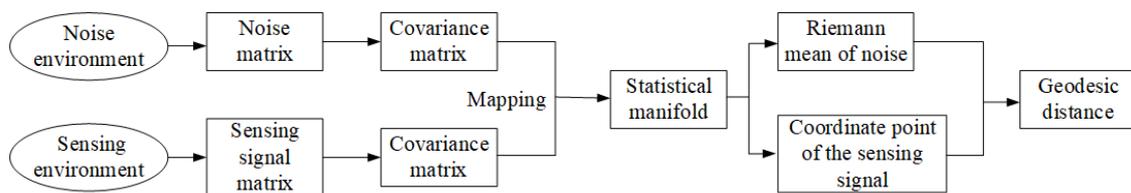

Figure 1. Spectrum sensing model.

In cognitive radio networks (CRNs), cooperative spectrum sensing (CSS) can not only overcome the physical limitations of single node spectrum sensing, but also achieve fast, accurate, and reliable sensing performance in the low SNR environments [22]. Cooperative spectrum sensing can be divided into three types: distributed cooperative spectrum sensing, centralized cooperative spectrum sensing, and relay cooperative spectrum sensing. The paper uses centralized spectrum sensing, as shown in Figure 2. The centralized spectrum sensing system has a fusion center (FC). Each SU does not need to exchange the sensing information, but sends the decision data to the FC. The FC performs a fusion decision on the received information according to the decision criteria.
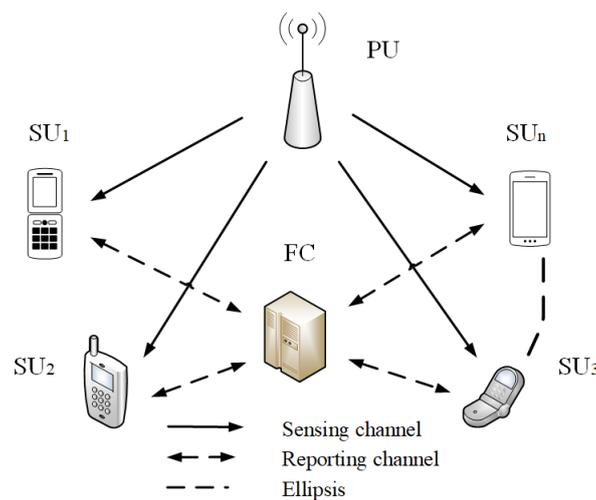

Figure 2. Centralized cooperative spectrum sensing. FC, fusion center.

Suppose that there are M secondary users in the CRNs and the number of sampling points per SU is N. H_0 indicates that the PU signal does not exist, and H_1 indicates that the PU signal exists. Therefore, the sensing signal under the two assumptions can be expressed as:

$$x_m(n) = \begin{cases} w_m(n), & H_0 \\ s_m(n) + w_m(n), & H_1 \end{cases}, \tag{1}$$

$m = 1, 2, \cdots, M$ and $n = 1, 2, \cdots, N$, where $s_m(n)$ represents the PU signal and $w_m(n)$ represents the Gaussian white noise signal. False alarm probability and detection probability can be defined as:

$$P_{fa} = P[H_1|H_0], \tag{2}$$

$$P_d = [H_1|H_1]. \tag{3}$$

The signal sensed by SU constitutes a vector matrix $X = [x_1, x_2, \ldots, x_M]^T$, where $x_m = [x_m(1), x_m(2), \ldots, x_m(N)]$ represents the signal sample value of the mth SU. Therefore, the dimension of matrix X is $M \times N$:

$$X = [x_1, x_2, \ldots, x_M]^T$$
$$= \begin{bmatrix} x_1(1) & x_1(2) & \cdots & x_1(N) \\ x_2(1) & x_2(2) & \cdots & x_2(N) \\ \vdots & \vdots & \ddots & \vdots \\ x_M(1) & x_M(2) & \cdots & x_M(N) \end{bmatrix}. \tag{4}$$

Thus, the statistical covariance matrix of the sensing signal:

$$R = \frac{1}{N}XX^T = R_S + R_W, \tag{5}$$

where R_S and R_W represent the covariance matrix of the noise and signal, respectively. The probability density function given an γ-dimensional vector subject to a zero-mean Gaussian distribution is:

$$p(x|R) = \frac{1}{\sqrt{(2\pi)^\gamma \det R}} \exp(-\frac{1}{2}x^T R^{-1} x), \tag{6}$$

where R is the ϖ-dimensional covariance matrix. The parameter model composed of the parameterized probability distribution family can be expressed as $S = \{p(x|R)|R \in C^{\varpi \times \varpi}\}$, where $C^{\varpi \times \varpi}$ is an open set in the ϖ-dimensional vector space. In information geometry, the structure of S is a microscopic manifold, also called a statistical manifold, where R is the coordinate point on the statistical manifold. Then, both R_W and $R_S + R_W$ under the two hypotheses can map the sensing signals to the statistical manifolds as coordinates on the manifold.

3. Spectrum Sensing Based on Information Geometry

The mapping between the sampled signal and the manifold was clarified in the previous section. Next, the calculation method of the two point distance on the statistical manifold will be introduced. Here, the geodesic distance is used to measure the distance between two points to reflect the characteristics of the signal.

3.1. Statistical Manifold and Fisher Metric

The parameterized probability density of a statistical manifold in an information geometry can be expressed as [23]:

$$S = \{p(x; \xi)|\xi \in \Theta \subset C^{\varpi \times \varpi}\}, \tag{7}$$

where the vector in the parameter space obeys a multivariate Gaussian distribution with a mean of zero. In $\xi = (\mu, \sigma)^T$, μ is the mean and σ^2 is the variance. In the literature [24], S is regarded as a statistical manifold with a dimension of ϖ. Each probability density $p(x; \xi)$ is a point in S where the coordinates are ξ. Therefore, the covariance matrix of the sensing signal in spectrum sensing can be mapped to the manifold. The difference between the sensing signals can be measured by the distance between

the two points on the statistical manifold. For the statistical manifold S, the Fisher information matrix plays the role of the Riemann metric, called the Fisher metric. Given the point $p(x;\xi) \in S$, the Fisher metric $G(\xi) = [g_{ab}(\xi)]$ at P on the manifold S is:

$$g_{ab}(\xi) = E_\xi[\frac{\partial}{\partial \xi^a} \ln p(x;\xi) \frac{\partial}{\partial \xi^b} \ln p(x;\xi)], \tag{8}$$

where E_ξ represents the mathematical expectation of $p(x;\xi)$. Therefore, we calculate the distance between two points on the statistical manifold according to Equation (8).

3.2. Geodesic Distance and Riemann Mean

On the manifold S, assume that the arbitrary curve between the two points θ_1 and θ_2 is $\theta(t)$, $t_1 \le t \le t_2$, $\theta(t_1) = \theta_1$, $\theta(t_2) = \theta_2$. Thus, the distance between θ_1 and θ_2 on curve $\theta(t)$ is [25]:

$$D(\theta_1, \theta_2) = \int_{t_1}^{t_2} \sqrt{(\frac{d\theta}{dt})^T G(\theta)(\frac{d\theta}{dt})} dt, \tag{9}$$

where $G(\theta)$ represents $G(\xi) = [g_{ab}(\xi)]$ in Equation (8). It can be seen from the above equation that the distance between θ_1 and θ_2 is related to curve $\theta(t)$. When the minimum distance is obtained by Equation (9), the curve is called the geodesic line, and the minimum distance is called the geodesic distance. The signal sensed by SU is a multivariate Gaussian distribution family with the same mean, but different covariance matrices. Therefore, for two arbitrary points Q_1 and Q_2 on S, we can calculate the geodesic distance between Q_1 and Q_2 [26]:

$$\begin{aligned} D(Q_1, Q_2) &= \| \log(Q_1^{-1/2} Q_2 Q_1^{-1/2}) \|_F \\ &= \sqrt{\Sigma_{\nu=1}^n \log^2(\lambda_\nu)}, \end{aligned} \tag{10}$$

where λ_ν is the νth eigenvalue of $Q_1^{-1/2} Q_2 Q_1^{-1/2}$.

For the L points $Q_l (l = 1, 2, \ldots, L)$ in the statistical manifold, when L is two, the Riemann mean $\overline{Q}$ is located at the midpoint connecting the Q_1 and Q_2 geodesics, whose value is [27]:

$$\overline{Q} = Q_1^{1/2}(Q_1^{-1/2} Q_2 Q_1^{-1/2})^{1/2} Q_1^{1/2}. \tag{11}$$

For the case of $L > 2$, the Riemann mean $\overline{Q}$ is difficult to find directly. According to the gradient descent algorithm in [28,29], the iterative calculation method of $\overline{Q}$ can be derived, so the expression of the Riemann mean can be obtained as:

$$\overline{Q}_{\alpha+1} = \overline{Q}_\alpha^{1/2} \exp(\frac{\tau}{L} \Sigma_{l=1}^L \log(\overline{Q}_\alpha^{-1/2} Q_l \overline{Q}_\alpha^{-1/2})) \overline{Q}_\alpha^{1/2}, \tag{12}$$

where τ is the iteration step size with $0 \le \tau \le 1$, α is the number of iteration steps, and l is the ordinal number of the point.

We can calculate the Riemann mean of the noise in the noisy environment by Equation (12). Then, we calculate the geodesic distance of the signal mapped to the manifold to the Riemann mean according to Equation (10). Finally, the geodesic distance is the input data for the neural network.

4. Spectrum Sensing Based on the Deep Neural Network

The whole training process is shown in Figure 3. The SU is greatly affected by SNR when performing signal detection. Low SNR will result in low detection performance of the system. Therefore, the input data of deep neural network are the detection result of multiple SUs under different SNRs.
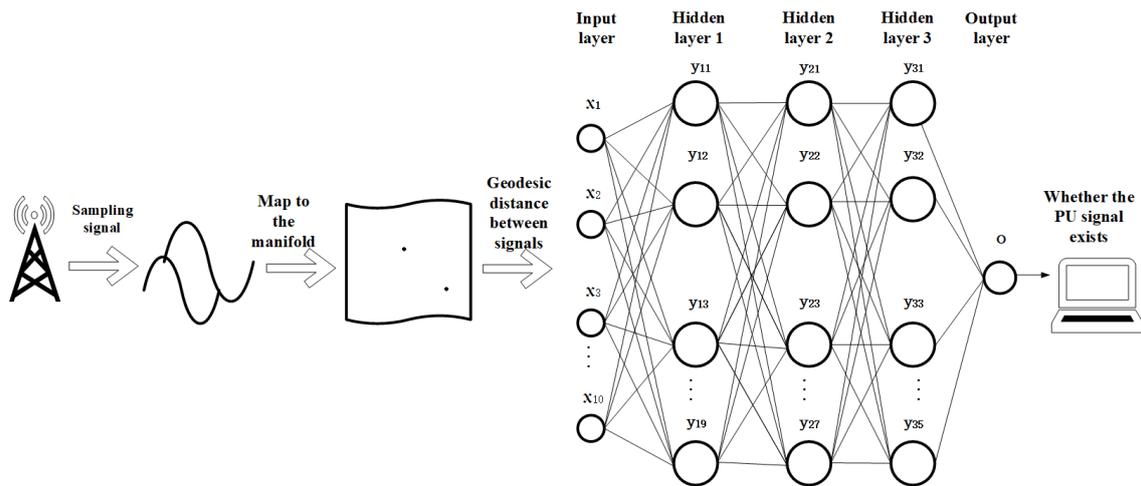
Figure 3. The learning model of deep neural network.

4.1. Determination of the Dataset

First, multiple SUs sense the energy value of only the presence of noise and the signal doped with noise under different SNR conditions. Then, with the same SNR, the geodesic distance between the noise and the sensing signal is calculated. Finally, the dataset composed of the geodesic distance is used as the input data of deep neural network. This paper divides the input data into three parts: training dataset, validation dataset, and test dataset. The training dataset is used to train the neural network. The validation dataset is used to select the hyperparameters of the neural network. The test dataset is used to test the classification ability of the network.

4.2. Selection of Hyperparameters in the Deep Neural Network

The selected structure of deep neural network is composed of an input layer, an output layer, and three hidden layers.

Input layer: Take the geodesic distance between the signals extracted by the information geometry method as input data. We select the sensing signals at 10 different SNRs; thus, the input layer selects 10 neurons.

Output layer: The output of deep neural network is to determine whether the PU signal exists, so we set the output layer as a neuron.

Hidden layer: We use a neural network with three hidden layers. In SNR = -18 dB~SNR = -12 dB, two SUs sense 10 times every 0.01 SNR. Then, the geodesic distance between the sensing signals is calculated at the SNR. The input data thus obtained are 1200 sets, and each set of features is 11. The 11th column represents the classification label. The first 1000 groups in the input data are used as training data; 1001~1100 groups are used as verification data; and 1101~1200 groups are used as test data. Table 1 was obtained by simulation using the validation dataset. As shown in Table 1, by comparing the accuracy of different hidden layers, it is shown that the network with three hidden layers has the highest accuracy.

Table 1. Comparison of testing accuracy using different hidden layers.

Number of Experiments	1	2	3	4	5	6	7	8	9	10	Average Value
1 hidden layer (%)	88	92	89	92	89	93	94	90	92	87	90.6
2 hidden layer (%)	89	93	93	90	92	92	92	89	95	91	91.6
3 hidden layer (%)	92	93	89	94	94	91	93	91	92	93	92.2
4 hidden layer (%)	94	94	91	93	89	91	91	92	9	93	91.9

Similarly, the number of neurons in the three hidden layers in the depth model was verified with the verification data, which were 9, 7, and 5, respectively. The experimental results are shown in Table 2.

Table 2. Comparison of testing accuracy using different numbers of neurons.

Number of Experiments	1	2	3	4	5	6	7	8	9	10	Average Value
5/2/2 (%)	94	90	92	89	92	93	90	94	90	89	91.3
9/7/5 (%)	92	94	89	94	96	89	90	89	94	93	92.0
20/10/8 (%)	88	95	91	93	93	87	92	91	92	92	91.4

As shown in Table 3, by comparing the two activation functions of Sigmoid and Tanh, the activation function of this model is set as the Tanh function.

Table 3. Comparison of testing accuracy using different activation functions.

Number of Experiments	1	2	3	4	5	6	7	8	9	10	Average Value
Sigmoid(%)	95	91	89	92	92	88	95	88	91	89	91.0
Tanh(%)	95	95	96	90	94	93	88	94	92	89	92.6

Given two probability distributions p and q, the cross-entropy can be defined as:

$$H(p,q) = -\Sigma_{i=1}^{n} p(c_i) \cdot \log(q(c_i)), \tag{13}$$

where $p(c_i)$ is the true distribution for the ith data, $q(c_i)$ is the theoretical distribution of the neural network, and n is the number of samples.

4.3. Training of the Deep Neural Network

In this paper, the geodesic distance between signals is used as the input data. The training sample set is $[c_1, c_2, \ldots, c_z]$, and z is the number of training samples. Any one of the samples $c_f = [c_{f1}, c_{f2}, \ldots, c_{fW}]$, and W is the number of features. The geodesic distance is between the sensing signal and the noise Riemann mean. If the PU is present, set the target output value to one; else, set it to zero for training. Then, the target output value is $[d_1, d_2, \ldots, d_z]$, and any one of the target output values d is one or zero. The actual output value is $[0_1, 0_2, \ldots, 0_z]$, and the expected error is e_0. The error value after the kth iteration satisfies $e(k) < e_0$, $e(k) = |d(k) - o(k)|$.

4.4. Testing of the Deep Neural Network

After the network is trained, the test sample set $[c_{s+1}, c_{s+2}, \ldots, c_Z]$ is placed as input data in deep neural network for training. The actual output of the test sample set is $[o_{s+1}, o_{s+2}, \ldots, o_Z]$. Finally, it is judged whether the PU signal exists according to whether the data value is close to zero or one.

4.5. The Training Process of the Feedforward Neural Network

The learning rate is set to v, and the weight value between the previous layer neuron i and the current layer neuron j in the network model is w_{ij}.

When the network model is training forward, the input values of the hidden layer and the output layer are:

$$I_j = \Sigma_i w_{ij} O_i + \theta_j, \tag{14}$$

where O_i means the output value of neurons i from the previous layer. θ_j represents the bias term of neurons from the current layer. The output of neurons from the current layer is related to the activation function. Therefore, the output value of the neuron from the layer is:

$$O_j = \frac{1}{1 + e^{-I_j}}.$$
(15)

When the network model is training in the opposite direction, the value of the error for the output layer:

$$e_j = O_j(1 - O_j)(d_j - O_j),$$
(16)

where d_j is the label value of the neuron, and the value of the error for the hidden layer:

$$e_j = O_j(1 - O_j)\Sigma_h e_h w_{hj},$$
(17)

where w_{hj} is the weight of the neuron h from the higher hidden layer to the neuron j from the current hidden layer. e_h stands for the error of the neuron h from the higher hidden layer.

The training process of Feedforward neural network is shown in Algorithm 1.

Algorithm 1: Classification learning algorithm.

Initialization:	Weight values and bias terms for deep neural networks.
Step 1:	**While** $e \geq e_0$ do
Step 2:	**for** Training any sample c in sample set $[c_1, c_2, \ldots, c_z]$ do
	Forward training
Step 3:	**for** Each neuron j in the hidden or output layer, do
Step 4:	$I_j = \Sigma_i w_{ij}O_i + \theta_j, O_j = \frac{1}{1+e^{-I_j}}$
Step 5:	**End for**
	Reverse training
Step 6:	**for** Each neuron j in the output layer, do
Step 7:	$e_j = O_j(1 - O_j)(d_j - O_j)$
Step 8:	**End for**
Step 9:	**for** In the training process from the last hidden layer to the first hidden layer, each neuron j of the hidden layer, do
Step 10:	$e_j = O_j(1 - O_j)\Sigma_h e_h w_{hj}$
Step 11:	**End for**
Step 12:	**for** Every deviation θ_j in the network model, do
Step 13:	$\Delta\theta_j = (l)e_j, \theta_j = \theta_j + \Delta\theta_j$
Step 14:	**End for**
Step 15:	**End for**
Step 16:	**End while**

---

5. Simulation Experiment and Result Analysis

The spectrum sensing method combining information geometry and deep neural network was simulated by MATLAB. There are multiple signals in the actual environment, so this paper sets up a multi-signal scenario to verify the performance of the proposed algorithm.

5.1. Multi-Component Signal

The PU signal was a multi-component signal $S(t) = cos(t) + cos(4t + 0.2t^2)$, and the noise was ideal Gaussian white noise.

The number of SUs had a large impact on the sensing effect. In order to verify the influence of the number M of SUs on the detection performance, the number M of SUs was set to 2, 3, and 5 in the range of SNR = -21 dB$\sim$SNR = -15 dB. Similarly, we trained the first 1000 groups of input data as training data and then used the latter 200 samples as test data to test the network performance. In Figure 4, Pd represents the detection probability, and Pfa represents the false alarm probability. It can be seen from the figure that when the false alarm probability is the same, as the number M of SU increases, the sensing performance of the method also steadily increases.
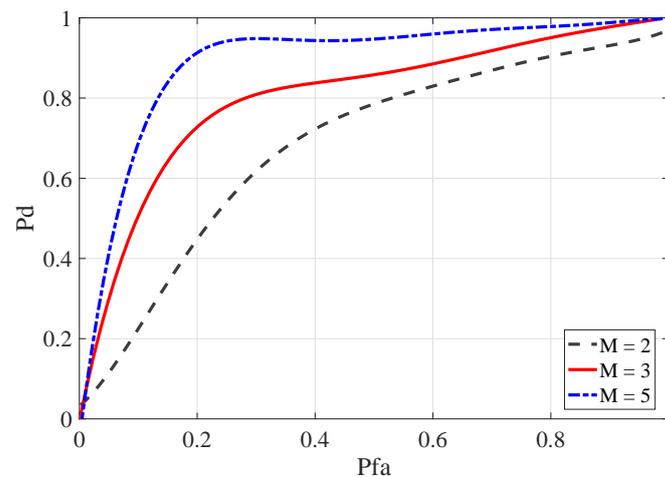
Figure 4. ROC curves for the number of SUs in different cases.

SNR was also one of the factors affecting the sensing performance of the algorithm. In order to explore the influence of SNR on the proposed algorithm, SNR was set to different value intervals, and the SUs were two. The same method was used to simulate the data extraction. It can be seen from Figure 5 that where the false alarm probability is the same, the sensing signal performance of the neural network is better in the interval where SNR is larger.
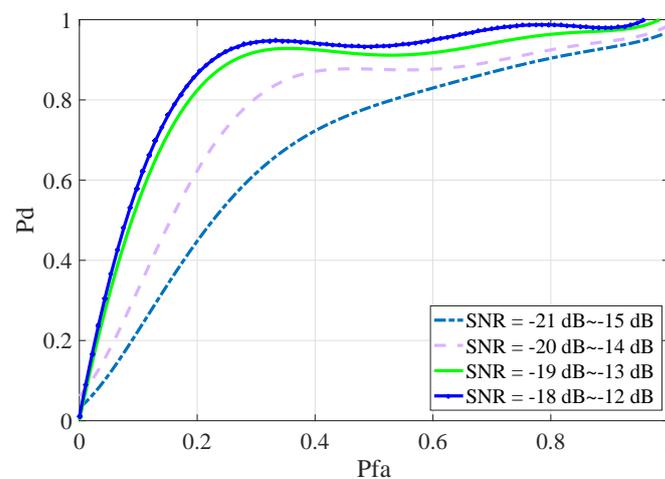
Figure 5. ROC curves at different SNRs.

In order to verify the superiority of the information geometry and spectrum sensing method, the following comparisons were made under the same SNR conditions, as shown in Figure 6. In SNR = −20 dB~SNR = −14 dB, IG-DNN, MME-DNN, and DMM-DNN calculated the sampling signal characteristics under the SNR, respectively. The MME determined the sensing signal by setting the decision threshold at a fixed SNR. Therefore, SNR = −14 dB, and the number of SUs was two for the simulation experiment. According to the ROC curves of MME and MME-DNN, it can be seen that when the false alarm probability were the same, the neural network algorithm had higher detection probability than the traditional classification algorithm. By comparing the IG-DNN algorithm with the other three methods, the application of information geometry to spectrum sensing improved the sensing performance. Inaccurate selection of threshold values can affect sensing performance, but deep neural network avoided setting decision thresholds.
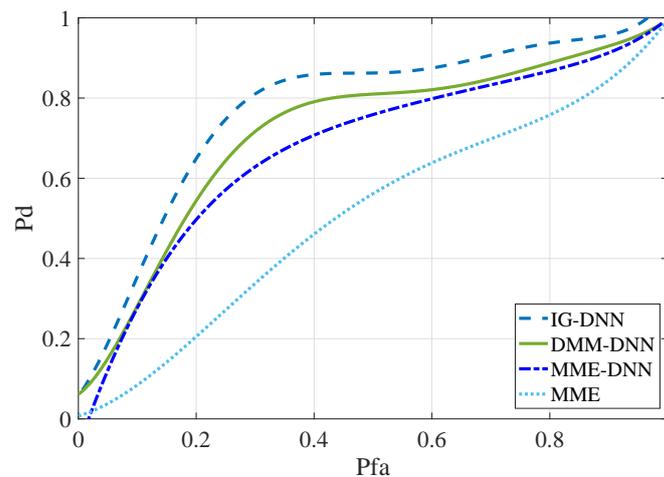
Figure 6. ROC curves for different methods with the same SNR (*a*)

Figure 7 compares the sensing algorithm proposed in this paper with other sensing algorithms. The IG-FCM algorithm was proposed in [17], and the MME-Kmeans algorithm was proposed in [30]. In the figure, IG-DNN is in the range of SNR = −20 dB∼SNR = −14 dB, and the other two methods are SNR = −17 dB for the simulation experiments. The number of SUs for the aforementioned methods was two. It can be observed from the ROC curve that the IG-DNN algorithm had the best sensing performance when the false alarm probability was the same.
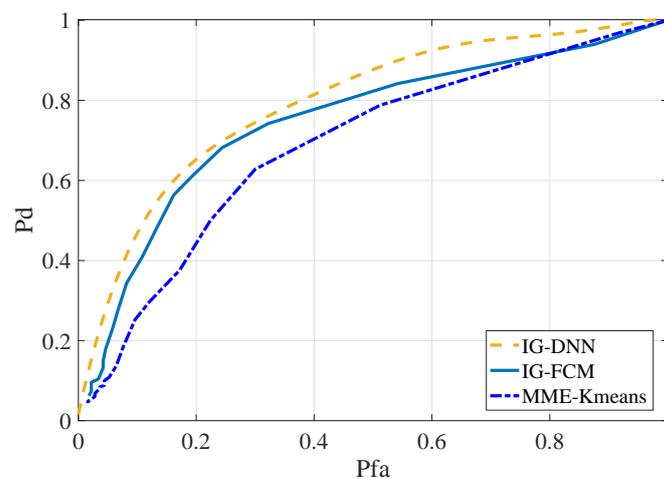
Figure 7. ROC curves for different methods with the same SNR (*b*).

5.2. AM Signal

The PU signal of the previous part was a multi-component signal, where the PU signal was replaced by an AM signal to further verify the IG-DNN algorithm. The AM signal was $S(t) = (1 + \sin(2 \cdot pi \cdot 2000t)) + \cos(2 \cdot pi \cdot fc \cdot t)$, where $t = n/fs$, *fs* is the sampling frequency, and *fc* is the signal frequency.

The IG-DNN algorithm was in the range of SNR = −18 dB∼SNR = −12 dB and SNR = −16 dB∼SNR = −10 dB, respectively. The other two methods were SNR = −15 dB and SNR = −13 dB, respectively. The number of SUs for the aforementioned methods was two. The simulation results are shown in Figures 8 and 9. According to the simulation results of the two graphs, the performance of the proposed IG-DNN algorithm was better than the other two methods.
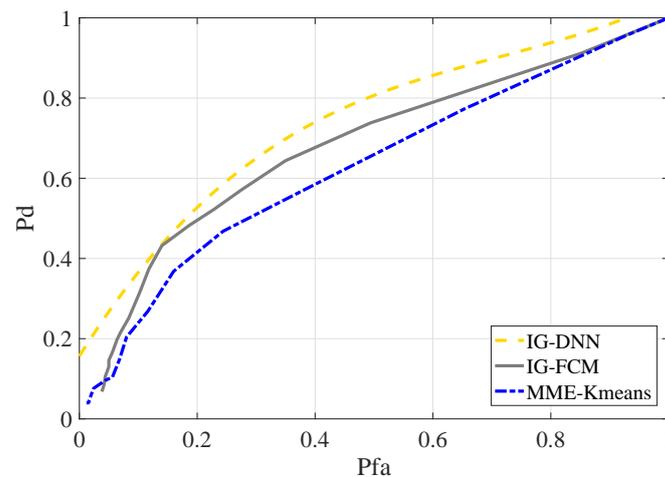
Figure 8. ROC curves for different methods in the AM signal scenario, SNR = −15 dB.
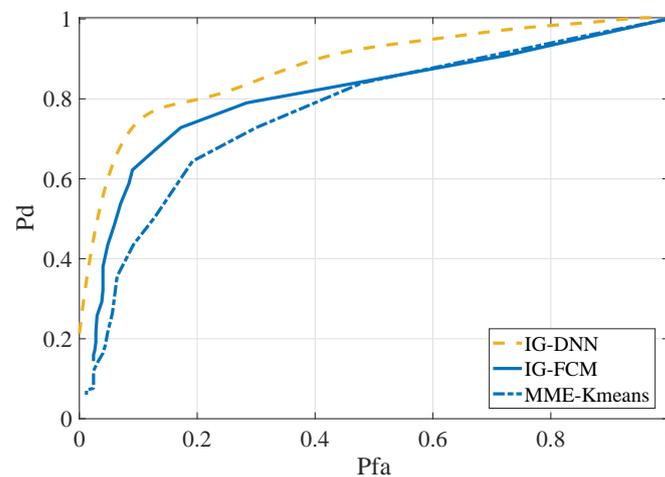

Figure 9. ROC curves for different methods in the AM signal scenario, SNR = −13 dB.

6. Conclusions

The IG-DNN algorithm proposed in this paper improved the sensing performance of the spectrum to some extent. The characteristics of the sensing signal were reflected in the statistical manifold, and the difference of the points on the manifold was calculated by the method of information geometry. Then, deep neural network was used to classify the geodesic distance to obtain the spectrum sensing classifier. Finally, the classifier was used to classify the unknown signals to determine whether the PU existed. Simulation experiments showed that the IG-DNN algorithm had better spectrum sensing performance. In future work, we are going to study array signal algorithms and a single sensor model. We will apply array signal algorithms or the single sensor model of spectrum sensing to further improve sensing performance.

References

1. Wang, Y.; Zhang, S.; Zhang, Y.; Wan, P.; Wang, S. A cooperative spectrum sensing method based on signal decomposition and K-medoids algorithm. *Int. J. Sens. Netw.* **2019**, *29*, 171–180. [CrossRef]
2. Wang, Y.; Ye, Z.; Wan, P.; Zhao, J. A survey of dynamic spectrum allocation based on reinforcement learning algorithms in cognitive radio networks. *Artif. Intell. Rev.* **2019**, *51*, 493–506. [CrossRef]
3. Wang, H.; Yang, E.H.; Zhao, Z.; Zhang, W. Spectrum sensing in cognitive radio using goodness of fit testing. *IEEE Trans. Wirel. Commun.* **2009**, *8*, 5427–5430. [CrossRef]
4. Zhang, L.; Xia, S. A new cooperative spectrum sensing algorithm for cognitive radio networks. In Proceedings of the 2009 ISECS International Colloquium on Computing, Communication, Control, and Management, Sanya, China, 8–9 August 2009; pp. 107–110.
5. Zheng, X.; Wang, J.; Wu, Q.; Chen, J. Cooperative spectrum sensing algorithm based on Dempster-Shafer theory. In Proceedings of the 2008 11th IEEE Singapore International Conference on Communication Systems, Guangzhou, China, 19–21 November 2008; pp. 218–221.
6. Cao, K.T.; Yang, Z. DET Cooperative Spectrum Sensing Algorithm Based on Random Matrix Theory. *J. Electron. Inf. Technol.* **2010**, *1*, 129–134. [CrossRef]
7. Zeng, Y.; Koh, C.L.; Liang, Y.C. Maximum eigenvalue detection: Theory and application. In Proceedings of the 2008 IEEE International Conference on Communications, Beijing, China, 19–23 May 2008; pp. 4160–4164.
8. Cao, K.T.; Yang, Z. A novel cooperative spectrum sensing algorithm based on random matrix theory. In Proceedings of the 2010 6th International Conference on Wireless Communications Networking and Mobile Computing, Chengdu, China, 23–25 September 2010; pp. 1–4.
9. Zeng, Y.; Liang, Y.C. Eigenvalue-based spectrum sensing algorithms for cognitive radio. *IEEE Trans. Commun.* **2009**, *57*, 1784–1793. [CrossRef]
10. Wang, Y.X.; Lu, G.Y. DMM based spectrum sensing method for cognitive radio systems. *J. Electron. Inf. Technol.* **2010**, *32*, 2571–2575. [CrossRef]
11. Yao, D.; Liu H.J. Support vector machine based spectrum sensing algorithm in cognitive radio. *Electron. Des. Eng.* **2018**, *21*, 1–5.
12. Lei, K.J.; Tan, Y.H.; Yang, X.; Wang, H.R. A K-means clustering based blind multiband spectrum sensing algorithm for cognitive radio. *J. Cent. South Univ.* **2018**, *25*, 2451–2461. [CrossRef]
13. Amari, S.I.; Nagaoka, H. *Methods of Information Geometry*; American Mathematical Society: Providence, RI, USA, 2007; pp. 1044–1068.
14. Wang, Y.; Chen, Q.; Li, J.; Wan, P.; Pang, S. A Spectrum Sensing Algorithm Based on Information Geometry and K-medoids Clustering. In Proceedings of the International Conference on Cloud Computing and Security, Haikou, China, 8–10 June 2018; pp. 219-230.
15. Chen, Q.; Wan, P.; Wang, Y.; Li, J.; Xiao, Y. Research on cognitive radio spectrum sensing method based on information geometry. In Proceedings of the International Conference on Cloud Computing and Security, Nanjing, China, 16–18 June 2017; pp. 554–564.
16. Wang, Y.; Zhang, S.; Zhang, Y.; Wan, P.; Li, J.; Li, N. A Cooperative Spectrum Sensing Method Based on Empirical Mode Decomposition and Information Geometry in Complex Electromagnetic Environment. *Complexity* **2019**, *2019*, 5470974. [CrossRef]
17. Zhang, S.; Wang, Y.; Li, J.; Wan, P.; Zhang, Y.; Li, N. A cooperative spectrum sensing method based on information geometry and fuzzy c-means clustering algorithm. *EURASIP J. Wirel. Commun. Netw* **2019**, *17*, doi:10.1186/s13638-019-1338-z. [CrossRef]
18. Ding, Q.; Zou, W.; Zhou, Z.; Li, B.; Ye, Y. A blind spectrum-sensing method based on Bartlett decomposition. In Proceedings of the 2011 6th International ICST Conference on Communications and Networking in China (CHINACOM), Harbin, China, 17–19 August 2011; pp. 639–644.
19. Xie, J.; Liu, C.; Liang, Y.C.; Fang, J. Activity Pattern Aware Spectrum Sensing: A CNN-Based Deep Learning Approach. *IEEE Commun. Lett.* **2019**, *23*, 1025–1028. [CrossRef]

20. Ke, D.; Huang, Z.; Wang, X.; Li, X. Blind Detection Techniques for Non-Cooperative Communication Signals Based on Deep Learning. *IEEE Access* **2019**, *7*, 89218–89225. [CrossRef]

21. Liu, Y.H.; Luo, S.W.; Li, A.J.; Huang, H.; Wen, J.W. Information geometry on extendable hierarchical large scale neural network model. In Proceedings of the 2003 International Conference on Machine Learning and Cybernetics, Xi'an, China, 2–5 November 2003; pp. 1380–1384.

22. Khan, A.A.; Rehmani, M.H.; Reisslein, M. Cognitive radio for smart grids: Survey of architectures, spectrum sensing mechanisms, and networking protocols. *IEEE Commun. Surv. Tutor.* **2015**, *18*, 860–898. [CrossRef]

23. Wang, M.; Ning, Z.H.; Xiao, C.; Li, T. Sentiment classification based on information geometry and deep belief networks. *IEEE Access* **2018**, *6*, 35206–35213. [CrossRef]

24. Calin, O.; Udrişte, C. *Geometric Modeling in Probability and Statistics*; Springer: Berlin, Germany, 2014; pp. 3–49.

25. Menendez, M.L.; Morales, D.; Pardo, L.; Salicru, M. Statistical tests based on geodesic distances. *Appl. Math. Lett.* **1995**, *8*, 65–69. [CrossRef]

26. Menendez, M.L.; Morales, D.; Pardo, L.; Salicru, M. A distance between multivariate normal distributions based in an embedding into the Siegel group. *J. Multivar. Anal.* **1990**, *8*, 223–242.

27. Balaji, B.; Barbaresco, F.; Decurninge, A. Information geometry and estimation of Toeplitz covariance matrices. In Proceedings of the International Radar Conference, Lille, France, 13–17 October 2014; pp. 1–4.

28. Moakher, M. A differential geometric approach to the geometric mean of symmetric positive-definite matrices. *SIAM J. Matrix Anal. Appl.* **2005**, *26*, 735–747. [CrossRef]

29. Lenglet, C.; Rousson, M.; Deriche, R.; Faugeras, O. Statistics on the manifold of multivariate normal distributions: Theory and application to diffusion tensor MRI processing. *J. Math. Imaging Vis.* **2006**, *25*, 423–444. [CrossRef]

30. Zhang, Y.; Wan, P.; Zhang, S.; Wang, Y.; Li, N. A spectrum sensing method based on signal feature and clustering algorithm in cognitive wireless multimedia sensor networks. *Adv. Multimed.* **2017**, *2017*, 2895680. [CrossRef]