An Anomaly Detector Using Filtering Stockwell Transform and Siamese Convolutional Neural Network in Radio Monitoring (AnoFSTSCNN)

Xiao Yan Wang, Rong Rong Qian, Chong Xing Huang, Ming Huang, and Jing Jing Yang

Abstract – Detecting anomalous behavior in the radio spectrum is a demanding task due to high levels of interference behaviors caused by massive wireless devices. Stockwell transform (ST) has grown in popularity for the analysis of nonstationary and nonlinear signals, but it is yet to be adequately explored in the radio monitoring domain. An approach that consists of FST (filtering Stockwell transform) and SCNN (siamese convolutional neural network) for radio spectrum anomaly detection (AnoFSTSCNN) is proposed in this paper. Four types of anomaly behaviors, including tone, chirp, pulse, and noise frequency modulation (FM) are simulated using MATLAB. Simulation results show that the AnoFSTSCNN has a 13.18% and 29.52% improvement in detection performance over short-time Fourier transform (STFT) and ST, respectively. The proposed approach provides a possible reference solution for radio monitoring and anomalous signal detection.

1. Introduction

A large number of wireless devices are applied to emerging applications to accommodate the massive traffic volume, leading to various high levels of interference behaviors [1, 2]. Considering the increasingly complex wireless channel environment, it is significant to design an adaptive and efficient detector to identify the source of harmful interference in these circumstances.

Although the Fourier transform (FT) spectral parameters are useful for detecting radio anomalies [3], the Fourier domain does not exhibit any time domain characteristics, leading to suboptimal feature extraction. Short-time Fourier transform (STFT) does not yield a multiresolution analysis due to the fixed bandwidth for signal decomposition at all frequencies [4, 5]. With an underlying assumption of a stationary signal, wavelet analysis is very useful [6]. More recently, Stockwell transform (ST) techniques for the analysis of nonstationary and nonlinear signals [7] have been widely applied in signal processing [8]. Despite popularity with nonstationary signal analysis, it is yet to be adequately explored in radio interference monitoring. Motivated by the excellent performance of ST, we extract higher order features from Stockwell images to detect and locate anomalies in the radio spectrum.

In our previous work [9], we used generative adversarial networks based on massive sample learning to detect anomalies in the radio spectrum. However, a large quantity of training samples is required, while storage and computing resources are greatly consumed. Recently, a Siamese convolutional neural network (SCNN) based on similarity calculation performs well with very limited samples applied to few-shot learning tasks and anomaly detection [10, 11]. This work presents a more effective anomaly detection method based on filtering ST (FST) and SCNN (AnoFSTSCNN). It was verified on a simulated interference data set containing the following four types of common jamming signals: 1) tone anomaly, continuous interference consisting of one or more sinusoids with random center frequency; 2) pulse anomaly, time-pulsed signal with random starting time and duration; 3) chirp anomaly, pulsed chirp signal with random center frequency and hopping rate; and 4) noise frequency modulation (FM) anomaly, noise FM signal obtained by Gaussian white noise through FM. The AnoFSTSCNN learns from limited samples and provides accurate detection and better application prospects. We organize the rest of the article as follows. Section 2 describes the proposed method for anomaly detection. The simulation signal model and data processing are illustrated in Section 3, and conclusions are drawn in Section 4.

2. Proposed Method

The four interference signals used in the simulation are tone, chirp, pulse, and FM. Tone is a continuous interference consisting of one or more sinusoids with random center frequency. Chirp indicates the pulsed chirp signal with random center frequency and hopping rate. Pulse means the time-pulsed signal with random starting time and duration. FM denotes the signal of which frequency changes with modulated noise. Continuous phase frequency shift keying (CPFSK)

Manuscript received 10 December 2022.

Xiao Yan Wang, Rong Rong Qian, Ming Huang, and Jing Jing Yang are with the School of Information Science and Engineering, Yunnan University, No.2 North Cuihu Road, Kunming, 650091, China; e-mail: xiaoyan wang2020@163.com, r.qiang@ynu.edu.cn, huangming@ynu.edu.cn, yangjingjing@ynu.edu.cn.

Chong Xing Huang is with Faculty of Social and Historical Sciences, University College London, Gower Street, London, WC1E 6BT, United Kingdom; e-mail: chongxing.huang.21@ucl.ac.uk.

This work was supported by the National Natural Science Foundation of China (grants 62261059, 61963037, and 61863035) and the 14th Postgraduate Scientific Research Innovation Foundation of Yunnan University (grant KC22222553).

modulation has the advantages of high bandwidth use, concentration of the main flap of the spectrum, and fast decay of the side flap roll-off. Signal and interference are expressed by (1).

In (1), f_c, f_d, and ϕ_0 denote the carrier frequency, peak frequency deviation, and random initial phase, respectively. $v(t)$ expresses the modulated signal. P_J and N_J present the jamming power and the number of tone, respectively. f_i is the frequency of the i_{th} jamming tone, and ϕ_i is the phase difference between the i_{th} component and the carrier of the hopping frequency slot. f_J is the starting frequency (at time $t=0$), Δf is the sweep rate of the chirp signal, and θ_J denotes initial phase of chirp. n_1 and n_2 are the period of the starting and ending times of the pulse interference, respectively, and w_t is a zero-mean Gaussian variable. U_j,w_j, K_{FM}, $u(t')$, and φ denote signal amplitude, constant, FM rate, modulated noise, and initial phase of the FM interference, respectively.

$$\begin{cases} \text{Signal} : s(t) = A\cos[2\pi f_c + 4\pi f_d \int_{-\infty}^{t} v(\tau)d\tau + \phi_0] \\ \text{Tone} : J_t(t) = \sqrt{2P_J/N_J}\sum_{i=1}^{N_J}\cos(2\pi f_i + \phi_i) \\ \text{Pulse} : J_P = \begin{cases} w_t, t \in [n_1, n_2] \\ 0, \text{else} \end{cases} \\ \text{Chirp} : J_c(t) = \sqrt{2P_J}\cos[2\pi(f_J + 0.5\Delta f t^2) + \theta_J] \\ \text{FM} : J_{\text{FM}}(t) = U_j\cos[w_j t + 2\pi K_{\text{FM}}\int_0^t u(t')dt' + \phi] \end{cases} \tag{1}$$

We use interference-to-signal ratio (ISR) to represent the different signal strengths of the interference signal to evaluate the antijamming performance, which is defined as $\text{ISR} = 10\log 10(P_J/P_S)$, where P_S denotes the signal power.

The primary task of anomaly detection is to decide whether the received signal is normal or abnormal. The discrimination between the two hypotheses can be modeled as

$$\begin{aligned} H_0 &: h(t) = s(t) + w(t) \\ H_1 &: h(t) = s(t) + j(t) + w(t) \end{aligned} \tag{2}$$

where $h(t)$ is the received signal, $j(t)$ represents an unexpected signal generated by either malfunctioning equipment or malicious jammers, $s(t)$ is a primary user's signal, $w(t)$ is additive white Gaussian noise, and t is the index of discrete time slots.

Because the temporal variation is the most important feature in the spectrum usage, STFT is used to obtain the time–frequency representation (TFR) of the received signal as

$$\begin{cases} \text{STFT}(\tau, f) = \int_{-\infty}^{+\infty} h(t)g(t-\tau)e^{-i2\pi ft}dt \\ g(t) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{t^2}{2\sigma^2}} \end{cases} \tag{3}$$

where the window function is $g(t)$ and σ is the variance that decides the width of the window function. The width of the window is a constant during time–frequency analysis (by convolution) at any point (τ, f) from (3), which means the insufficient precision and poor frequency resolution for a narrow window or a low time resolution for a wide window. In fact, we expect to follow these rules to reach the best TFR: 1) use a narrow window function in analysis of high-frequency components; and 2) use a wider window function in analysis of low-frequency components. To balance the temporal and frequency resolution, an adjustment factor δ is added to make the window adaptive, which can be written as

$$\text{STFT}'(\delta, \tau, f) = \int_{-\infty}^{+\infty} h(t)\frac{1}{\sqrt{2\pi}\delta}e^{-i2\pi ft}dt \tag{4}$$

To realize the adaptive adjustment of the window function according to the frequency distribution of the signal, the window length is defined as $\delta_f = 1/|f|$. Therefore, the expression of ST, whose window width inversely varies with the signal frequency in the continuous domain, is

$$S(\tau, f) = \int_{-\infty}^{+\infty} h(t)\left[\frac{|f|}{\sqrt{2\pi}}e^{-\frac{f^2(\tau-t)^2}{2}}e^{-i2\pi ft}\right]dt \tag{5}$$

The window function of ST can be written as $g'(t) = \frac{|f|}{\sqrt{2\pi}}e^{-\frac{t^2 f^2}{2}}$, indicating that the window function of ST satisfies the previous two requirements of a good TFR. In particular, the linearity of the ST makes it possible to analyze signals with additional noise [7]:

$$S\{h(t)\} = S\{s(t)\} + S\{j(t)\} + S\{w(t)\} \tag{6}$$

where $S\{h(t)\}$ is ST matrix of received signal $h(t)$, $S\{s(t)\}$ and $S\{w(t)\}$ indicate the constant terms, and $S\{j(t)\} = S\{h(t)\} - S\{s(t)\} - S\{w(t)\}$ is the abnormal term that is expected to be detected. Therefore, $S\{s(t)\} + S\{w(t)\}$ can be marked as redundant information, which can be filtered out from $S\{h(t)\}$. This is the core idea of FST as presented in this article.

The Siamese network architecture ties two identical neural networks, with shared weights, by a loss function. The identical neural network takes a data set $\chi = \{x_n \in \mathbb{R}^D\}_{n=1}^N$, where N and D denote the samples and feature space dimension as input and maps all vectors to a new space $\chi' = \{x'_n \in \mathbb{R}^{D'}\}_{n=1}^N$ to preserve small pairwise distances between the high-dimensional and low-dimensional vectors by minimizing the contrastive loss [11], typically $D' \ll D$. Contrastive loss is defined as the distance between two identical neural networks' output. For the given pairs $x_1, x_2 \in R^D$, the squared Euclidean distance is described as $d(x_1, x_2) = \|G(x_1) - G(x_2)\|$, the binary label y equals 1 (x_1, x_2 is similar) or 0 (x_1, x_2 is dissimilar), and the contrastive loss takes the form as

$$\begin{aligned} \ell(y, x_1, x_2) = &\frac{1}{2}(1-y)\max\{0, 1 - d(x_1, x_2)\} \\ &+ \frac{1}{2}yd(x_1, x_2) \end{aligned} \tag{7}$$

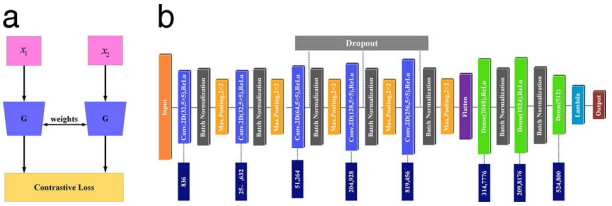

Figure 1. Network structure of the proposed AnoFSTSCNN: (a) Siamese network architecture; and (b) basic convolutional neural networks G in AnoFSTSCNN.

A designed SCNN structure is depicted in Figure 1a, which describes the overall architecture, and Figure 1b details the identical neural network.

3. Simulation and Results

In this section, a communication system based on CPFSK modulation and demodulation is simulated by using MATLAB 2021a communication tools (version X) to generate the normal signal. In this system, the original binary code stream signal is modulated by CPFSK and passed through a Gaussian channel, and tone interference, pulse interference, chirp interference, and noise FM interference is added, respectively. Both the normal and abnormal signals are simulated and preprocessed based on FST, ST, and STFT, respectively, to obtain the data set. The tested SNR (signal-to-noise ratio) is in the range of −10 dB to 10 dB, with a 5 dB interval, and the ISR is in the range of −12 dB to 20 dB, with a 4 dB interval.

Figure 2 describes the processing of generating a Stockwell image. First, the initial input vector A is sent to ST to generate a complex time–frequency distribution matrix S, and the redundant information (obtained by computing the mean value of the ST matrix of the 100 normal sample images) is filtered out from S to yield more details S'. Finally, the energy matrix C is calculated by performing $abs(\cdot)$ to S'. Example Stockwell images are shown in Figure 3. Tone, chirp, pulse, and FM interference are located at the bottom, center, left, and top of the image, respectively. Figure 3a is from the original ST algorithm, which contains not only the anomalous part but also redundant information that may mask the abnormal part in the low ISR environment. The proposed method appropriately filters redundant information, as shown in Figure 3b, where the redundant information in the red box is filtered out, while the information in the region where the anomaly is located is well preserved.

The Stockwell image pairs are marked as 1 for similar pairs and 0 for dissimilar pairs through SCNN. Under the supervised learning paradigm, the SCNN

Figure 2. Block diagram of FST-based preprocessing method for CPFSK signal vectors.

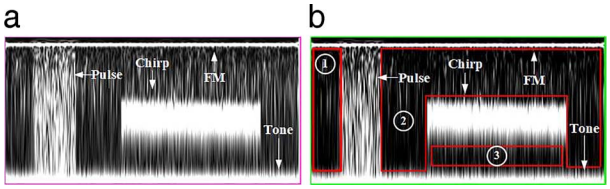

Figure 3. Examples of Stockwell images: (a) Stockwell image obtained by DST; and (b) Stockwell image obtained by proposed method.

maximizes the representation of different labels and minimizes the representation of the same label, which means that the goal of the training process is to make the distance between similar image pairs as close as possible and the distance between dissimilar images as far as possible. This work mainly uses accuracy, receiver operating characteristic (ROC) curve, and the area under ROC curve as the evaluation metrics.

Figure 4 depicts the performance of the detecting model. Figure 4a describes that the model is more sensitive to tone at lower interference levels and less sensitive to the remaining three types of anomalies. The detection accuracy of the model increases at a higher interference level, and the detection probability exceeds 80% at ISR $>$ 4 dB. Figure 4b shows that the model easily detects tone interference under different channel conditions. The detection accuracy of the model increases at higher SNRs. Note that the AnoFSTSCNN still has a good result in the environment with low SNR. For instance, the detection probability is about 67% at SNR $= -10$ dB.

The ROC curves of different TFR algorithms and anomalies are illustrated in Figure 5. Lines in red, green, and blue represent the ROC curves on FST, ST, and FTST, respectively. Figures 5a–5d represent the performance of different TFR algorithms under the chirp, tone, pulse and FM, respectively. According to the ROC curves, the FST-based one is the best, and the STFT-based ROC curve is suboptimal. Anomaly detection ability can be greatly improved by filtering out redundant information from the original ST matrix.

The overall performance of the proposed method is evaluated when only the presence of exceptions is considered without concerning the anomaly type. Figure 6 reveals the overall detection accuracy using FST, ST, and STFT. The FST-based method has the best

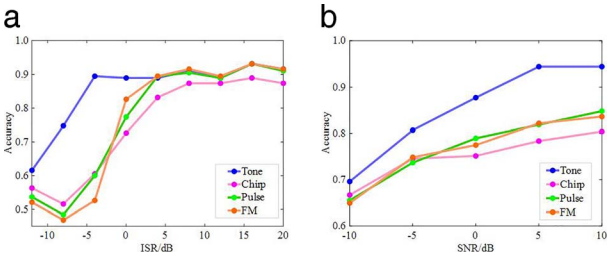

Figure 4. FST-based detection accuracy with different channel conditions and interference levels.

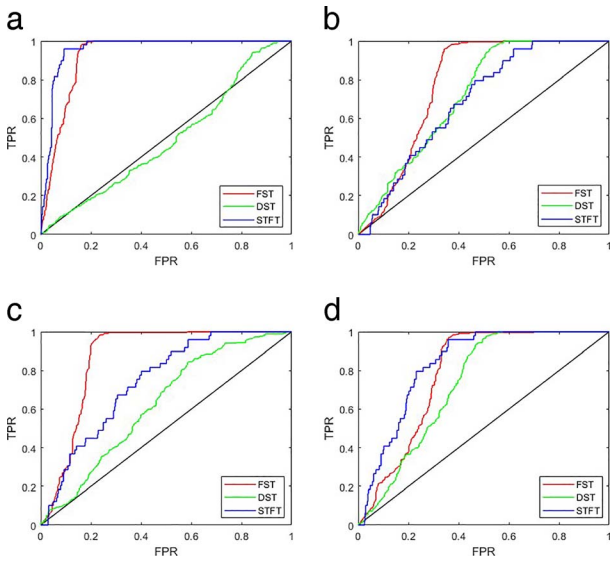

Figure 5. FST-, ST-, and STFT-based ROC curves with different anomalies. (a) Tone. (b) Chirp. (c) FM. (d) Pulse.

performance among the three TFR analysis methods. The average detection accuracy of the FST-, ST-, and STFT-based methods is 84.52%, 55%, and 71.34% at SNR = 5 dB, which reveals that the FST-based method outperforms the others.

Based on the previous analysis, the FST is based on the original ST, and the anomalous part that is not of concern for anomaly detection is filtered out from the original ST image. Therefore, FST is computationally identical compared with ST. On the other hand, this article introduces the advantage of few-shot learning by using SCNN, which requires fewer samples to obtain a large performance gain. Although the proposed method achieves good performance, it has limitations at low SNR and high ISR conditions, which is a common problem that needs to be continuously improved.

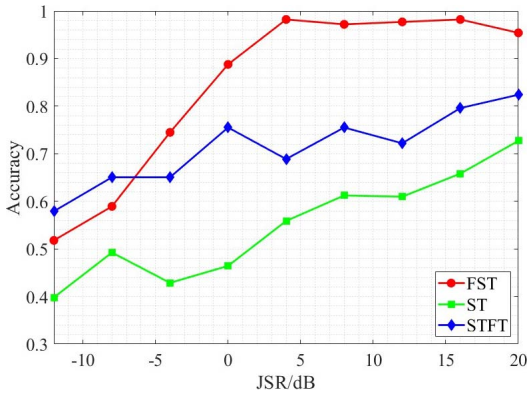

Figure 6. FST-, ST-, and STFT-based average accuracies with SNR = 5 dB.

4. Conclusion

In this article, we have proposed an approach AnoFSTSCNN for radio anomaly detection. The SCNN was designed to extract features from filtered Stockwell images and identify tone, chirp, FM, and pulse anomalies in the radio spectrum. The performance of the AnoFSTSCNN was verified on a simulation data set, and it was shown that the accuracy is far superior to the ST- and STFT-based methods, providing an alternative for radio monitoring and spectrum anomaly detection. In the future, we will apply the proposed method in a test network of radio monitoring.

5. References

1. Q. Lu, J. Yang, Z. Jin, D. Chen, and M. Huang, "State of the Art and Challenges of Radio Spectrum Monitoring in China," *Radio Science*, **52**, 10, October 2017, pp. 1261-1267.
2. M. Huang, D. Yang, D. Zhu, M. Yang, and J. Yang, "FM Broadcast Monitoring Using Artificial Intelligence," *Radio Science*, **55**, 4, February 2020, pp. 1-6.
3. J. Kowsik, T. B. Kumar, G. Mounika, D. V. Ratnam, and S. Raghunath, "Detection of Ionospheric Anomalies Based on FFT Averaging Ratio (FAR) Algorithm," International Conference on Innovations in Information, Embedded and Communication Systems, Coimbatore, India, March 19–20, 2015, pp. 1-3.
4. X. Zhou, J. Xiong, X. Zhang, X. Liu, and J. Wei, "A Radio Anomaly Detection Algorithm Based on Modified Generative Adversarial Network," *IEEE Wireless Communications Letters*, **10**, 7, April 2021, pp. 1552-1556.
5. G. Chen, G. Lu, and P. Yan, "Early Fault Detection of Rolling Element Bearings Based on Visibility Graph Modeling of Vibration Signals," International Conference on Prognostics and System Health Management, Jinan, China, October 23–25, 2020, pp. 198-203.
6. M. Golgowski and S. Osowski, "Anomaly Detection in ECG Using Wavelet Transformation," International Conference on Computational Problems of Electrical Engineering, Online Conference, Poland, September 16–19, 2020, pp. 1-4.
7. R. G. Stockwell, L. Mansinha, and R. Lowe, "Localization of the Complex Spectrum: The *S* Transform," *IEEE Transactions on Signal Processing*, **44**, 4, April 1996, pp. 998-1001.
8. D. Li, A. Ukil, K. Satpathi, and Y. M. Yeap, "Improved S Transform-Based Fault Detection Method in Voltage Source Converter Interfaced DC System," *IEEE Transactions on Industrial Electronics*, **68**, 6, April 2020, pp. 5024-5035.
9. X. Wang, J. Yang, L. Zhang, Q. Lu, and M. Huang, "Spectrum Monitoring of Radio Digital Video Broadcasting Based on an Improved Generative Adversarial Network," *Radio Science*, **56**, 8, July 2021, pp. 1-13.
10. X. Zhou, W. Liang, S. Shimizu, J. Ma, and Q. Jin, "Siamese Neural Network Based Few-Shot Learning for Anomaly Detection in Industrial Cyber-Physical Systems," *IEEE Transactions on Industrial Informatics*, **17**, 8, December 2020, pp. 5790-5798.
11. T. Li, Z. Hong, Q. Cai, L. Yu, Z. Wen and R. Yang, "BISSIAM: Bispectrum Siamese Network Based Contrastive Learning for UAV Anomaly Detection," in *IEEE Transactions on Knowledge and Data Engineering*, doi: 10.1109/TKDE.2021.3118727.