

Sequential Convolutional Recurrent Neural Networks for Fast Automatic Modulation Classification

KAISHENG LIAO¹, YAODONG ZHAO¹, JIE GU¹, YAPING ZHANG¹, AND YI ZHONG¹

Science and Technology on Electronic Information Control Laboratory, Chengdu 610000, China

Corresponding author: Yi Zhong (eric.yi.zhong@outlook.com)

ABSTRACT A novel and efficient end-to-end learning model for automatic modulation classification is proposed for wireless spectrum monitoring applications, which automatically learns from the time domain in-phase and quadrature data without requiring the design of hand-crafted expert features. With the intuition of convolutional layers with pooling serving as the role of front-end feature distillation and dimensionality reduction, sequential convolutional recurrent neural networks are developed to take complementary advantage of parallel computing capability of convolutional neural networks and temporal sensitivity of recurrent neural networks. Experimental results demonstrate that the proposed architecture delivers overall superior performance in signal to noise ratio range above -10 dB, and achieves significantly improved classification accuracy from 80% to 92.1% at high signal to noise ratio range, while drastically reduces the average training and prediction time by approximately 74% and 67%, respectively. Response patterns learned by the proposed architecture are visualized to better understand the physics of the model. Furthermore, a comparative study is performed to investigate the impacts of various sequential convolutional recurrent neural network structure settings on classification performance. A representative sequential convolutional recurrent neural network architecture with the two-layer convolutional neural network and subsequent two-layer long short-term memory neural network is developed to suggest the option for fast automatic modulation classification.

INDEX TERMS Automatic modulation classification, convolutional neural networks, cognitive radio, deep learning, recurrent neural networks, spectrum monitoring.

I. INTRODUCTION

Wireless spectrum monitoring over time, space and frequency is important for effective use of the scarce spectral resources in various commercial areas [1]–[5]. As an integral part of wireless spectrum monitoring systems, automatic modulation classification (AMC) is used to recognize modulation types without prior knowledge of the received signals and channel parameters [6]–[8]. AMC has been proven to be an essential capability for transmitter identification, wireless spectrum anomaly detection and radio environment awareness. It improves radio spectrum utilization and opens the possibility of intelligent decision for context-aware autonomous wireless spectrum monitoring systems.

The existing AMC approaches discussed in literature can be roughly brought down into the following

The associate editor coordinating the review of this manuscript and approving it for publication was Longxiang Gao¹.

two categories: (i) likelihood-based approaches; and (ii) feature-based approaches [9], [10]. For the first category, the likelihood-based approaches utilize hypothesis testing theory and form a judgment criterion by analyzing statistical characteristics of signals [11], [12]. In likelihood-based approaches, modulation classification is framed as Bayesian estimation to optimize the probability of classification. However, approaches of this type are not robust in the presence of unknown channel conditions and suffer from heavy computational load on their practical implementations. Traditional feature-based approaches mainly focus on expert feature extraction and classification criteria [13]–[18]. They utilize expert features such as higher order cyclic moments for modulation classification. It is easy and simple for these approaches to be implemented in practical systems. However, hand-crafting expert features and hard-coding rules for modulation classification make it difficult to scale to new modulation types in non-cooperative scenarios.

Recently, researchers in wireless communications have started to apply deep neural networks to cognitive radio tasks with some success [19]–[33]. The authors in [19], [24] demonstrated that convolutional neural networks (CNNs) trained on time domain in-phase and quadrature (IQ) data significantly outperform conventional expert feature-based approaches. The authors in [20], [23] utilized recurrent neural networks (RNNs) for learning temporal representations to achieve higher classification accuracy than that of the CNNs introduced in [19]. In [21], the authors directly adopted convolutional long short-term deep neural networks (CLDNNs) from voice processing domain. The authors in [34] developed a data-driven fusion method to obtain better classification accuracy using the combination of the two CNNs trained on different datasets. Ramjee *et al.* [32] performed a comparative study of various typical deep neural networks and reduced the training complexity by reducing the input dimensionality with subsampling techniques.

In autonomous wireless spectrum monitoring systems, online learning is fundamental for accommodating new emerging modulation types and complex environmental circumstances. Nevertheless, those RNN models delivering high classification accuracy suffer from computational complexity and long training time. In this work, we develop a novel and efficient sequential convolutional recurrent neural network (SCRNN) architecture combining parallel computing capability of CNNs with temporal sensitivity of RNNs. Experimental results demonstrate that our approach outperforms the state-of-the-art on classification performance, while significantly improves the rate of convergence compared with the CNN and RNN alone architectures.

The rest of the paper is organized as follows. In Section II, an overview of the modulation benchmark dataset is introduced, and the two baseline models are briefly explained. The proposed model and the parameters used for training along with other implementation details are clearly stated in Section III. Section IV details the classification results and discusses the advantages of the proposed model. Conclusions and future work are presented in Section V.

II. DATASET AND BASELINES

A. DATASET

In a wireless spectrum monitoring system, the received signal can be typically represented as:

$$r(t) = s(t) * h(t) + n(t) \quad (1)$$

where $s(t)$ denotes the noise free complex baseband envelope of the received signal, and $h(t)$ refers to the time varying impulse response of the transmitted wireless channel. $n(t)$ represents the additive white Gaussian noise (AWGN) reflecting thermal noise. The complex received signal $r(t)$ is commonly sampled in IQ format due to its simplicity.

A typical modulation dataset RadioML2016.10a generated by GNU Radio is used as the benchmark dataset for training

TABLE 1. Benchmark dataset parameters.

Dataset	RadioML2016.10a
Modulations	8 Digital Modulations: BPSK, QPSK, 8PSK, 16QAM, 64QAM, BFSK, CPFSK, and PAM4 3 Analog Modulations: WBFM, AM-SSB, and AM-DSB
Length per sample	128
Signal format	In-phase and quadrature (IQ)
Signal dimension	2×128 per Sample
Duration per sample	128 μ s
Sampling frequency	1 MHz
Samples per symbol	8
SNR Range	[-20 dB, -18 dB, -16 dB, . . . , 18 dB]
Total number of samples	220000 vectors
Number of training samples	198000 vectors
Number of test samples	22000 vectors

and evaluating the performance of the proposed architecture, similar as the MNIST dataset in the vision domain [35]. The dataset follows the signal representation as given in equation 1. Detailed parameter description of the dataset is shown in Table 1. Radio channel effects are relatively well characterized in the dataset. Channel imperfections such as multi-path fading, random walk drifting of carrier frequency oscillator and sample time clocks, AWGN, along with unknown scale, translation, and dilation transformation are introduced into the signal in the dataset for reflecting the real electromagnetic environment [35]. The dataset is labeled with both signal to noise ratio range (SNR) ground truth and modulation types.

B. BASELINES

The two models are chosen as the baselines for further comparisons due to their results showing the significant improvements upon expert feature-based approaches. Any further improvements should be considered state-of-the-art.

One is the CNN architecture proposed by O'shea *et al.* [19]. As shown in Fig. 1(a), the baseline model is a 4-layer network made up of two convolutional layers and two dense layers. Each hidden layer utilizes rectified linear unit (ReLU) activation functions and dropout of 50% except for a softmax activation function on the one-hot output layer. Adam optimizer and categorical cross entropy loss function are applied to the base model.

The other baseline model is proposed by Rajendran *et al.* [23], shown in Fig. 1(b). The model is comprised of two 128-unit long short-term memory (LSTM) layers and an 11-unit dense layer with a softmax activation. The first LSTM layer returns the full sequences while the second one just returns the last state. The dropout is also adopted to reduce overfitting. Adam optimizer and categorical cross entropy loss function are applied to the model. Note that this model learns from the time domain information of the modulation schemes using amplitude-phase format, instead of IQ format.

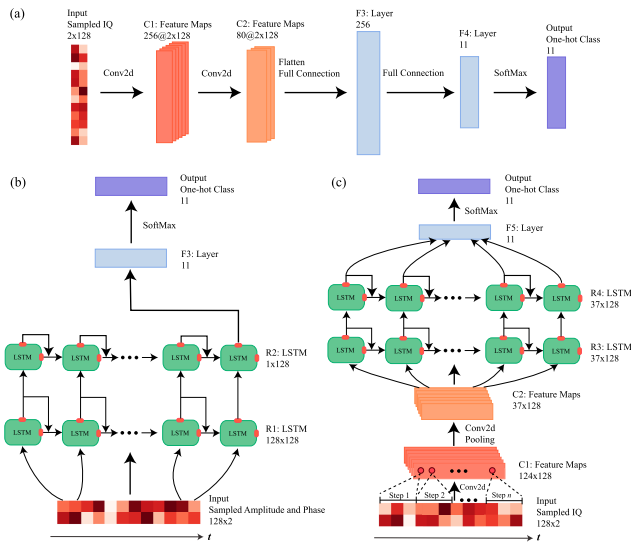


FIGURE 1. Schematic diagram of the (a) convolutional neural network (CNN) baseline model, (b) long short-term memory (LSTM) baseline model, and (c) proposed sequential convolutional recurrent neural network (SCRNN) model.

III. SEQUENTIAL CONVOLUTIONAL RECURRENT NEURAL NETWORKS

A. MOTIVATION

Generally, the received radio signals sampled at discrete time steps are of time domain sequences. In [23], a two-layer LSTM architecture is proposed and achieves a good classification accuracy of 86% at high SNRs. However, these models using RNNs suffer from much slower training time than that of the CNNs, due to their computational complexity and unparallel computing capability. Thus, a new novel and efficient SCRNN architecture is proposed with the combination of the speed and lightness of CNNs and the temporal sensitivity of RNNs. Furthermore, as a variant of RNN, LSTM is adopted instead of simple RNN in the proposed architecture to remember long-term dependencies and avoid the gradient vanishing problem. In SCRNN architectures, the convolutional layers with pooling acting as the role of front-end feature distillation and dimensionality reduction turn the long input sequences into much shorter representations of high-level features, which then become the input for subsequent LSTM layers to learn long-term temporal coherence of modulations.

B. MODEL DESCRIPTION

Fig. 1(c) provides the illustration of the proposed SCRNN architecture. As schematically shown in Fig. 1(c), the first and second convolutional layers each contain 128 5-tap filters except for the first one followed by a max-pooling layer with a pooling size of 3. The layer 3 and layer 4 are LSTM layers composed of 128 units each, and both return the full sequences. The last dense layer contains 11-class neurons representing the modulation schemes.

ReLU activation functions are applied to the convolutional and LSTM layers. The last dense layer utilizes a softmax

activation to achieve modulation classification. Dropout regularization combined with max norm has been proven to be of better performance for preventing overfitting. Categorical cross entropy is adopted as the loss function, which can be written as:

$$L = -\frac{1}{N} \sum_{i=1}^N \mathbf{y}_i \cdot \log(\hat{\mathbf{y}}_i) \quad (2)$$

where \mathbf{y}_i represents the ground truth in the form of one-hot encoding, and $\hat{\mathbf{y}}_i$ refers to the prediction. N denotes the training batch size. Adam optimizer with a learning rate of 0.001 is utilized due to its computational efficiency.

C. IMPLEMENTATION DETAILS

The total 220000 samples in the RadioML2016.10a dataset are split into two, one training set of 198000 (90%) samples and the other test set of 22000 (10%) samples. The dataset is split equally among all considered modulation types using the stratified sampling strategy. Instead of extracting the amplitude and phase features of the signals manually in advance [23], we adopted IQ components as input directly. A batch size of 128 is used on each training epoch and the early stop strategy is adopted.

All training and prediction are implemented in Keras library [36] on the backend of TensorFlow [37]. The Nvidia Cuda enabled Tesla K80 is used to speed up the calculation.

IV. RESULTS AND DISCUSSION

The classification performance of the models on the benchmark dataset is discussed in this section. We inspect and compare the classification accuracy and rate of convergence between the baseline models and the proposed SCRNN model. In addition, the varying kernel sizes, kernel types and layer depths are further investigated to find the optimal SCRNN architecture.

The classification accuracy of all the models are presented in Fig. 2. It can be seen that the proposed SCRNN model delivers a significantly improved accuracy of 92.1% at high SNRs. The CNN and LSTM model as baselines are compared to the proposed SCRNN model. It shows that the SCRNN model consistently achieves higher accuracy than the other

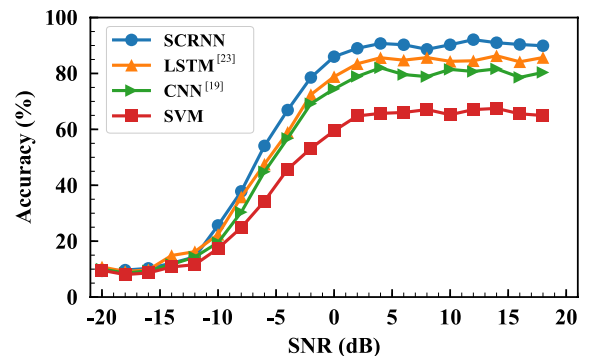


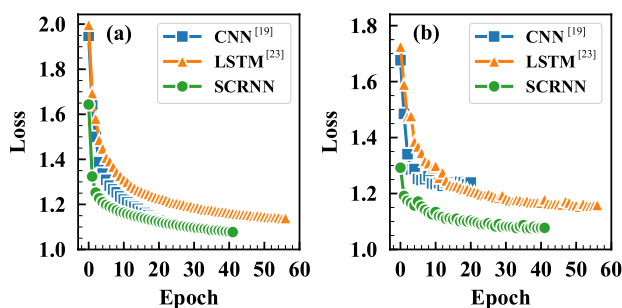
FIGURE 2. Classification accuracy comparison of the proposed SCRNN model with others on the benchmark dataset.

TABLE 2. Training and prediction time comparison between the two baseline models and the SCRNN model.

Models	CNN [19]	LSTM [23]	SCRNN
Number of trainable parameters	5369947	200075	398731
Training time per epoch (s)	30	800	280
Training epochs	20	56	41
Training time (s)	600	44800	11480
Prediction time ($\mu\text{s}/\text{sample}$)	1000	2000	661

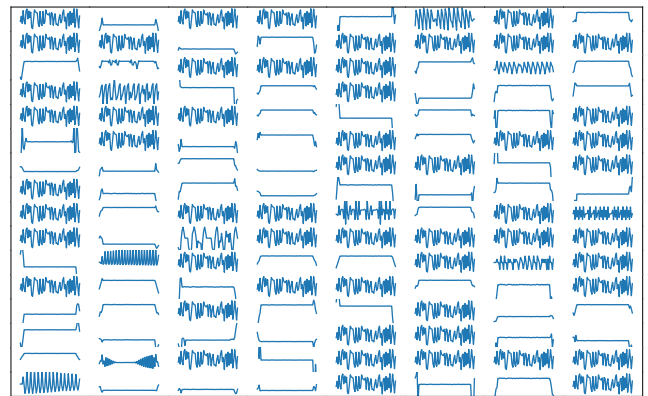
two baselines in the SNR range from -10 dB to 18 dB, and significantly outperforms the CNN baseline model by 12% and the LSTM baseline model by 6% improvement at high SNRs. Additionally, it is observed that the proposed SCRNN model achieves exceeding performance than that of the CNN and LSTM baseline models in the SNR range from -10 dB to 0 dB, where the two baseline models behave nearly the same. It implies that the convolutional layers of the SCRNNs serving as the role of feature distillation boost the learning ability of the temporal features under low SNR circumstances. The traditional support vector machine (SVM) approach showing relatively poor classification performance is also summarized in Fig. 2 for comparison. Note that all models are fed with the same training and test data of IQ format for this comparison except for the LSTM model with amplitude-phase format, and the standardization instead of L_2 normalization in [19] is adopted to scale the input for all models.

Fig. 3 shows the training history including the (a) training loss and (b) validation loss compared between the baseline models and the proposed SCRNN model. According to the training history, the LSTM baseline model achieves the second less loss value but remains the lowest rate of convergence; the CNN baseline model obtains faster rate of convergence but yields the largest loss value, while the proposed SCRNN model retains the fastest rate of convergence and achieves the least loss value among the three. The average training and prediction time together with the network size of the three models are compared in Table 2. It can be seen that though the introduction of the convolutional layers in the SCRNN leads to nearly double the network size, the average training and prediction time of the proposed SCRNN model are drastically reduced to only 280 seconds per epoch and 661 μs per sample

**FIGURE 3.** Training history including the (a) training loss and (b) validation loss between the baseline models and the proposed SCRNN model.

respectively, compared to 800 seconds per epoch and 2000 μs per sample of the LSTM model. These are fairly consistent with the insight that the convolutional layers with pooling before RNN serve as the role of feature distillation and dimensionality reduction, analogous to front-end matched filters, synchronizer and sampler for temporal features in typical wireless systems. Thus, the improved quality of the input for the SCRNN model makes it significantly reduce the training time and achieve the fastest prediction time.

To gain intuition on what convolution layers are learning in SCRNN architectures, the response patterns of the 128 filters learned by the first convolutional layer are illustrated in Fig. 4, showing that some filters encode expert-like patterns (i.e. BPSK-like pattern in row 1 column 6) and others even encode more complicated patterns. It further confirms that the convolutional layers of the SCRNNs act as the role of front-end feature distillation with coherent features refined and redundant features filtered out, enabling the improved rate of convergence.

**FIGURE 4.** Response patterns of the 128 filters learned by the first convolutional layer of the SCRNN.

To gain more insight into the SCRNN architecture, we further investigate the effects of various SCRNN structure settings varying CNN kernel sizes, CNN layer depths, CNN kernel numbers, RNN types and RNN layer depths on classification performance.

As shown in Fig. 5(a), varying the CNN kernel sizes of the SCRNN has minimal impact on classification performance. The architecture with kernel size of 5 produces slightly better classification accuracy than others in SNR range from 0 dB to 18 dB, while the architecture with kernel size of 3 leads to marginally higher classification accuracy in SNR range from -10 dB to -6 dB. The kernel size of 5 is used for the remaining experiments.

Increasing of the CNN layer depths with pooling reduces the input dimensionality for subsequent LSTM layers in the SCRNN architecture, and hence reduces the training time. Fig. 5(b) proves that the input dimensionality reduction shows very limited effects on classification performance. However, the performance of the LSTM baseline model starts to decay significantly when reducing the input dimensionality [32]. It is implied that the SCRNN architecture is much more

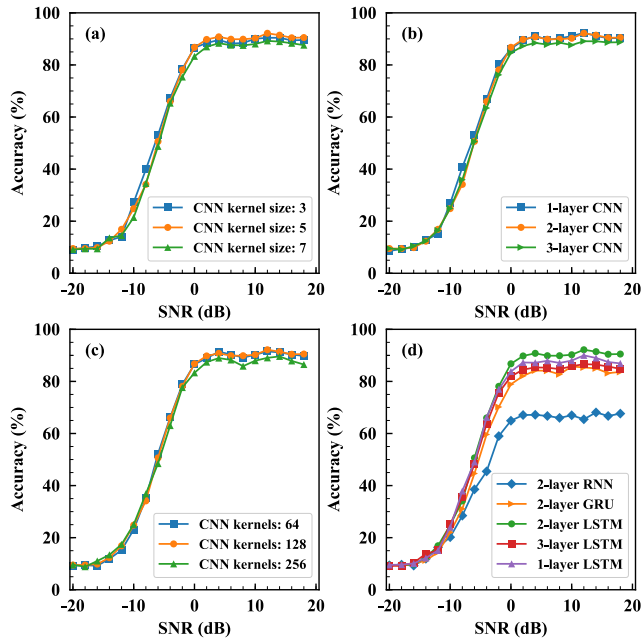


FIGURE 5. Classification performance of different SCRNN structure configurations for varying (a) CNN kernel sizes, (b) CNN layer depths, (c) CNN kernel numbers, (d) RNN types and layer depths.

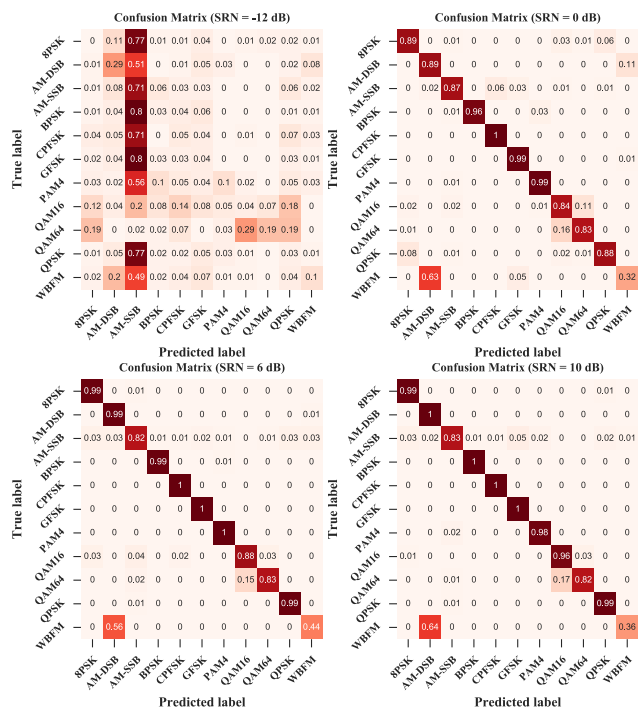


FIGURE 6. Confusion matrices for the optimized SCRNN architecture on the benchmark dataset at various SNRs.

robust to dimensionality reduction. Thus, it makes possible for deploying online learning model on autonomous wireless spectrum monitoring systems.

Fig. 5(c) provides that the 64-kernel and 128-kernel structures deliver the very similar performance, while the performance of 256-kernel structure starts to drop due to the overfitting. Fig. 5(d) shows the different settings of the RNN

types and layer depths in the SCRNN architecture. It can be observed that the performance of the LSTM type is apparently superior to that of the gated recurrent unit (GRU) and the simple RNN type. Experimental results of varying LSTM layer depths suggest that the 2-layer LSTM of the SCRNN achieves the best classification accuracy. Therefore, the optimal SCRNN architecture with the 2-layer CNN and subsequent 2-layer LSTM is recommended for online learning.

To evaluate how classification performance varies with SNRs, confusion matrices of the optimal SCRNN model at various SNRs are investigated. For a confusion matrix, each column represents the predicted modulation type and each row represents the real modulation type. The numerical value on each grid denotes the prediction probability of the corresponding modulation type.

As illustrated in Fig. 6, the diagonals become gradually sharper with increasing SNR, yet two primary confusions exist even at high SNRs. One is among the analog modulations. This is mainly due to the silent period existing in the analog audio signal [19]. The other is between QAM16 and QAM64 as the former is a subset of the latter.

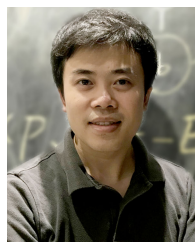
V. CONCLUSION

In this paper, a novel and efficient SCRNN architecture for AMC has been developed. Compared with the CNN and LSTM baseline models, the proposed architecture takes full advantage of the complementarity of CNNs and RNNs. Thus, it makes the classification accuracy deliver the state-of-the-art performance, improved from 80% to 92.1% at high SNRs. The average training and prediction time of the proposed architecture are significantly reduced by approximately 74% and 67% respectively, paving the way for deployment of online learning models on autonomous wireless spectrum monitoring systems. Response patterns learned by the proposed architecture have been investigated to better understand what feature pattern each filter in the convolutional layers is receptive to. Additionally, a comparative study of various structure settings of SCRNNs has been performed, and a representative SCRNN architecture with the 2-layer CNN and subsequent 2-layer LSTM was developed to recommend for fast AMC. Future work will focus on validation on radio signals with varying symbol rates and bandwidths. Second, unsupervised or deep reinforcement learning approaches for AMC should be investigated due to the lack of necessary signal labels in real wireless spectrum monitoring systems. Finally, stream learning without requiring to retrain the entire network from scratch is also a worthy direction for future research.

REFERENCES

- [1] M. Hoyhtya, A. Mammela, M. Eskola, M. Matinmikko, J. Kalliovaara, J. Ojaniemi, J. Suutala, R. Ekman, R. Bacchus, and D. Roberson, "Spectrum occupancy measurements: A survey and use of interference maps," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2386–2414, 4th Quart., 2016.
- [2] S. Barber, R. Petruschka, and E. de Castro, "Using wireless network access points for monitoring radio spectrum traffic and interference," U.S. Patent 10 430 731, Mar. 18 2004.

- [3] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 1, pp. 116–130, 1st Quart., 2009.
- [4] S. Zheng, S. Chen, L. Yang, J. Zhu, Z. Luo, J. Hu, and X. Yang, "Big data processing architecture for radio signals empowered by deep learning: Concept, experiment, applications and challenges," *IEEE Access*, vol. 6, pp. 55907–55922, 2018.
- [5] K. Thilina, K. W. Choi, N. Saquib, and E. Hossain, "Machine learning techniques for cooperative spectrum sensing in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 11, pp. 2209–2221, Nov. 2013.
- [6] C. Weber, M. Peter, and T. Felhauer, "Automatic modulation classification technique for radio monitoring," *Electron. Lett.*, vol. 51, no. 10, pp. 794–796, May 2015.
- [7] C. Clancy, J. Hecker, E. Stuntebeck, and T. O'Shea, "Applications of machine learning to cognitive radio networks," *IEEE Wireless Commun.*, vol. 14, no. 4, pp. 47–52, Aug. 2007.
- [8] M. Waqar Aslam, Z. Zhu, and A. Kumar Nandi, "Automatic modulation classification using combination of genetic programming and KNN," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2742–2750, Aug. 2012.
- [9] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Survey of automatic modulation classification techniques: Classical approaches and new trends," *IET Commun.*, vol. 1, no. 2, pp. 137–156, Apr. 2007.
- [10] A. K. Nandi and E. E. Azzouz, "Algorithms for automatic modulation recognition of communication signals," *IEEE Trans. Commun.*, vol. 46, no. 4, pp. 431–436, Apr. 1998.
- [11] J. L. Xu, W. Su, and M. Zhou, "Likelihood-ratio approaches to automatic modulation classification," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 41, no. 4, pp. 455–469, Jul. 2011.
- [12] W. Wei and J. M. Mendel, "Maximum-likelihood classification for digital amplitude-phase modulations," *IEEE Trans. Commun.*, vol. 48, no. 2, pp. 189–193, Feb. 2000.
- [13] B. Ramkumar, "Automatic modulation classification for cognitive radios using cyclic feature detection," *IEEE Circuits Syst. Mag.*, vol. 9, no. 2, pp. 27–45, Jun. 2009.
- [14] H.-C. Wu, M. Saquib, and Z. Yun, "Novel automatic modulation classification using cumulant features for communications via multipath channels," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 3098–3105, Aug. 2008.
- [15] C.-S. Park, J.-H. Choi, S.-P. Nah, W. Jang, and D. Y. Kim, "Automatic modulation recognition of digital signals using wavelet features and SVM," in *Proc. 10th Int. Conf. Adv. Commun. Technol.*, Feb. 2008, pp. 387–390.
- [16] A. Swami and B. M. Sadler, "Hierarchical digital modulation classification using cumulants," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 416–429, Mar. 2000.
- [17] S.-Z. Hsue and S. S. Soliman, "Automatic modulation classification using zero crossing," *IEE Proc. F Radar Signal Process.*, vol. 137, no. 6, pp. 459–464, Dec. 1990.
- [18] S. S. Soliman and S.-Z. Hsue, "Signal classification using statistical moments," *IEEE Trans. Commun.*, vol. 40, no. 5, pp. 908–916, May 1992.
- [19] T. J. O'Shea, J. Corgan, and T. C. Clancy, "Convolutional radio modulation recognition networks," in *Proc. Int. Conf. Eng. Appl. Neural Netw.* Cham, Switzerland: Springer, 2016, pp. 213–226.
- [20] D. Hong, Z. Zhang, and X. Xu, "Automatic modulation classification using recurrent neural networks," in *Proc. 3rd IEEE Int. Conf. Comput. Commun. (ICCC)*, Dec. 2017, pp. 695–700.
- [21] N. E. West and T. O'Shea, "Deep architectures for modulation recognition," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DySPAN)*, Mar. 2017, pp. 1–6.
- [22] A. Ali, F. Yangyu, and S. Liu, "Automatic modulation classification of digital modulation signals with stacked autoencoders," *Digit. Signal Process.*, vol. 71, pp. 108–116, Dec. 2017.
- [23] S. Rajendran, W. Meert, D. Giustiniano, V. Lenders, and S. Pollin, "Deep learning models for wireless signal classification with distributed low-cost spectrum sensors," *IEEE Trans. Cognit. Commun. Netw.*, vol. 4, no. 3, pp. 433–445, Sep. 2018.
- [24] T. J. O'Shea, T. Roy, and T. C. Clancy, "Over-the-air deep learning based radio signal classification," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 168–179, Feb. 2018.
- [25] M. Sadeghi and E. G. Larsson, "Adversarial attacks on deep-learning based radio signal classification," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 213–216, Feb. 2019.
- [26] D. Zhang, W. Ding, B. Zhang, C. Xie, H. Li, C. Liu, and J. Han, "Automatic modulation classification based on deep learning for unmanned aerial vehicles," *Sensors*, vol. 18, no. 3, p. 924, Mar. 2018.
- [27] C.-F. Teng, C.-C. Liao, C.-H. Chen, and A.-Y.-A. Wu, "Polar feature based deep architectures for automatic modulation classification considering channel fading," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Nov. 2018, pp. 554–558.
- [28] J. Sun, G. Wang, Z. Lin, S. G. Razul, and X. Lai, "Automatic modulation classification of cochannel signals using deep learning," in *Proc. IEEE 23rd Int. Conf. Digit. Signal Process. (DSP)*, Nov. 2018, pp. 1–5.
- [29] M. Kulin, T. Kazaz, I. Moerman, and E. De Poorter, "End-to-end learning from spectrum data: A deep learning approach for wireless signal identification in spectrum monitoring applications," *IEEE Access*, vol. 6, pp. 18484–18501, 2018.
- [30] B. Tang, Y. Tu, Z. Zhang, and Y. Lin, "Digital signal modulation classification with data augmentation using generative adversarial nets in cognitive radio networks," *IEEE Access*, vol. 6, pp. 15713–15722, 2018.
- [31] S. Zheng, P. Qi, S. Chen, and X. Yang, "Fusion methods for CNN-based automatic modulation classification," *IEEE Access*, vol. 7, pp. 66496–66504, 2019.
- [32] S. Ramjee, S. Ju, D. Yang, X. Liu, A. El Gamal, and Y. C. Eldar, "Fast deep learning for automatic modulation classification," 2019, *arXiv:1901.05850*. [Online]. Available: <http://arxiv.org/abs/1901.05850>
- [33] Z. Zhang, H. Luo, C. Wang, C. Gan, and Y. Xiang, "Automatic modulation classification using CNN-LSTM based dual-stream structure," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 13521–13531, Nov. 2020.
- [34] Y. Wang, M. Liu, J. Yang, and G. Gui, "Data-driven deep learning for automatic modulation recognition in cognitive radios," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 4074–4077, Apr. 2019.
- [35] T. J. O'shea and N. West, "Radio machine learning dataset generation with GNU radio," in *Proc. GNU Radio Conf.*, 2016, vol. 1, no. 1.
- [36] F. Chollet, "Keras: The Python deep learning library," *Astrophys. Source Code Library*, Jun. 2018, Art. no. ascl:1806.022.
- [37] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, and M. Kudlur, "Tensorflow: A system for large-scale machine learning," in *Proc. 12th USENIX Symp. Operating Syst. Design Implement. (OSDI)*, 2016, pp. 265–283.



KAISHENG LIAO was born in Sichuan, China, in 1987. He received the B.S. degree from Sichuan University, in 2010, and the Ph.D. degree from the Shanghai Institute of Technical Physics, Chinese Academy of Sciences, in 2015. He is currently an Associate Researcher with the Science and Technology on Electronic Information Control Laboratory, Chengdu, China. His main research interests include spectrum sensing, compressed sensing, and machine learning algorithms for radio signal processing.



His main research interest includes cognitive and adaptive signal processing.

YAODONG ZHAO was born in Sichuan, China, in 1986. He received the B.S. degree in communication engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2008, and the Ph.D. degree in signal and information processing from the University of Chinese Academy of Sciences, Beijing, China, in 2013. He is currently an Associate Researcher with the Science and Technology on Electronic Information Control Laboratory, Chengdu, China.



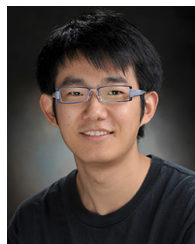
and cognitive radio signal processing.

JIE GU was born in Sichuan, China, in 1974. He received the B.S. degree in information engineering and the Ph.D. degree in signal and information processing from the University of Electronic Science and Technology of China, Chengdu, China, in 1997 and 2004, respectively. He is currently a Researcher with the Science and Technology on Electronic Information Control Laboratory, Chengdu, China. His current research interests include adaptive array signal processing,



YAPING ZHANG was born in Mianyang, Sichuan, China, in 1992. He received the B.S. and M.S. degrees in control engineering from the Nanjing University of Aeronautics and Astronautics, in 2014 and 2017, respectively. Since 2017, he has been working as an Associate System Engineer with the Science and Technology on Electronic Information Control Laboratory, Chengdu, China. His current work involves the optimal design of unmanned system, path planning and machine

learning algorithm for signal processing.



YI ZHONG was born in Chengdu, Sichuan, China, in 1987. He received the B.S. degree from Hong Kong Baptist University and the M.S. degree from Simon Fraser University. He is currently an Associate Researcher with the Science and Technology on Electronic Information Control Laboratory, Chengdu, China. His main research interests include spectrum sensing, swarm intelligence, and machine learning algorithms for radio signal processing.

• • •