IEICE TRANS. ELECTRON., VOL.E106–C, NO.8 AUGUST 2023

BRIEF PAPER

EMRNet: Efficient Modulation Recognition Networks for Continuous-Wave Radar Signals

Kuiyu CHEN[†a)], **Jingyi ZHANG**[†], **Shuning ZHANG**[†], **Si CHEN**[†], ***Nonmembers*, *and* Yue MA**[†], ***Member***

SUMMARY Automatic modulation recognition(AMR) of radar signals is a currently active area, especially in electronic reconnaissance, where systems need to quickly identify the intercepted signal and formulate corresponding interference measures on computationally limited platforms. However, previous methods generally have high computational complexity and considerable network parameters, making the system unable to detect the signal timely in resource-constrained environments. This letter firstly proposes an efficient modulation recognition network(EMRNet) with tiny and low latency models to match the requirements for mobile reconnaissance equipments. One-dimensional residual depthwise separable convolutions block(1D-RDSB) with an adaptive size of receptive fields is developed in EMRNet to replace the traditional convolution block. With 1D-RDSB, EMRNet achieves a high classification accuracy and dramatically reduces computation cost and network paraments. The experiment results show that EMRNet can achieve higher precision than existing 2D-CNN methods, while the computational cost and parament amount of EMRNet are reduced by about 13.93× and 80.88×, respectively.

key words: *automatic modulation recognition, radar signals, efficient, low latency, adaptive size of receptive fields*

1. Introduction

Automatic modulation recognition (AMR) of radar signals has attracted considerable attention in military and civil applications [1], [2]. In the past decades, there have been two main categories of AMR methods for radar signals. In the first category, handcrafted features have been extracted from radar signals to classify waveforms [3], [4]. However, the selection of handcrafted features dramatically depends on the experience of researchers. Moreover, these features have to be re-selected when the system needs to recognize new waveforms. Thus, alternative deep learning technologies are proposed to overcome the shortcomings of handcrafted features. Deep learning, simulating the human brain for learning the inner laws and representation levels of sample data, can automate the feature-extracting process. [5] utilized entire time-frequency images (TFI) as inputs for LeNet-5 [6] to identify radar signals. In [7], TFIs denoised by singular value decomposition are presented to a shrinking ResNet-50 [8], which enhances the recognition performance. [9], [10] adopted similar convolutional neural networks (CNN) to recognize radar signals from their two-dimensional (2D) transform domain. Benefits from the powerful learning ability of CNN, the recognition accuracy of radar signals have been further improved.

For most existing research, original 1D radar signals have to be mapped into a 2D transform domain [11], [12]. And then, 2D CNN is employed to extract features from 2D transform-domain images. However, these approaches suffer several limitations. First, the process of 2D transformation, such as time-frequency transformation, is time-consuming and complicated. Second, 2D CNNs mainly focus on images with a large number of samples. Thus, the training and optimization of 2D CNNs will cost too much time. Moreover, these methods concentrated on pulse-wave (PW) signals and depended on an unrealistic assumption that the pulse width of signals has been accurately obtained. Continuous-wave (CW) radar has better anti-jamming capabilities and has been widely applied in target tracking and short-range detection system. Compared with PW radar, CW radar spreads the signal energy over a much longer time interval, resulting in a dramatic increase in the computational complexity of subsequent processing. Thus, an efficient recognition method for CW radar signals is urgent to be proposed.

Recently, 1D CNN has been developed to deal with 1D micro-Doppler signals for human activity classification and demonstrates outstanding results in terms of precision and complexity [13]. 1D CNNs are deployed to capture high-level representations by utilizing a set of 1D kernels. Therefore, the sequence can be directly input to 1D CNN, which avoids the complex 2D transformation. The autocorrelation function (ACF) calculates the similarity of the signal with its time lag, which can keep the computational complexity at a low level. Thus, it is an interesting issue to combine 1D CNN with ACF to identify CW radar signals. Nevertheless, a hurdle hampering the application of 1D CNN is that numerous network weights generated by high-dimensional input signals are difficult to be optimized. Additionally, the fixed receptive field of the traditional framework limits the learning ability of networks [14].

In this letter, we propose a novel efficient modulation recognition network (EMRNet) to overcome these existing hurdles in CNN and then utilize it to recognize CW radars. The main contributions of this study are threefold:

• 1D residual depthwise separable convolutions block (1D-RDSB) is firstly developed to extract features from autocorrelation signals. Compared with the standard 1D convolutions block, 1D-RDSB has a more efficient architecture, which dramatically reduces computation costs while main-

Manuscript received June 3, 2022.
Manuscript revised October 19, 2022.
Manuscript publicized March 24, 2023.

[†]The authors are with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China.
a) E-mail: kuiyu_chen@163.com (Corresponding author)

DOI: 10.1587/transele.2022ECS6006

taining accuracy.

• A multi-scale features fusion incorporating a lightweight attention mechanism is proposed for the adaptive aggregation of information from different receptive fields, further improving the representational capability of the network.

• The sufficiently tiny model of EMRNet is specifically tailored for mobile and resource-constrained environments, which can be directly stored on almost all FPGA.

2. Signal Model and Preprocessing

2.1 Signal Model

The received radio frequency (RF) radar signal is down-converted and sampled discretely as below:

$$y[n] = r[n] + a[n] \tag{1}$$

where $r[n]$ symbolizes the noise-free radar signal, $a[n]$ denotes channel noise supposed to be the additive white Gaussian noise (AWGN).

This research considers ten types of CW radar signals. There are frequency modulated signals: linear frequency modulation (LFM) and sinusoidal frequency modulation (SFM); phase modulated signals: binary phase-shift keying (BPSK) and Frank code signal; digital frequency modulated signals: frequency-shift keying (FSK) and 4FSK; amplitude modulated signals: binary amplitude-shift keying (BASK); no modulated signal (NS); and combined modulation signals: LFM-BPSK and SFM-BPSK. Detailed definitions of these CW radar signals can be found in related publications [3]–[5].

2.2 Autocorrelation Feature

Autocorrelation analysis is a nonlinear tool for radar signal processing, which measures the similarity between signals separated by different time lags. The autocorrelation function (ACF) can detect non-randomness in signal, which can be computed as:

$$\begin{aligned} A[m] &= \sum_n y[n] \times y^*[n-m] \\ &= \sum_n [x[n] + w[n]] \times [x^*[n-m] + w^*[n-m]] \\ &= \sum_n [x[n] \times x^*[n-m]] + \delta[m] \end{aligned} \tag{2}$$

Under the background of AWGN, the cross-correlation between signal and noise is approximately equal to zero. ACF of noise can equal to $\delta[m]$. Furthermore, $A[0]$ can be replaced by $A[1]$ to eliminate the impact of noise.

3. Methodology

The network architecture of the proposed EMRNet is illustrated in Fig. 1. Due to the symmetry of ACF, half of the autocorrelation sequence is selected as the input signal. At the beginning of the EMRNet, the input layer has the same size as the received 1D signal sequence. Then, seven 1D-RDSBs connected in series are applied to extract abstract information from ACF. The output of each block serves as the input of the next block. The reason for utilizing seven 1D-RDSBs is that we make a trade-off between recognition performance and the number of operations. Features learned by seven 1D-RDSBs are flattened into a 1D vector and then mapped to sample label space by fully connected layers (FC).

1D-RDSB, as the critical component of EMRNet, is designed to decrease computational complexity and paraments needed while maintaining excellent precision. The structure of 1D-RDSB will be discussed in detail below.

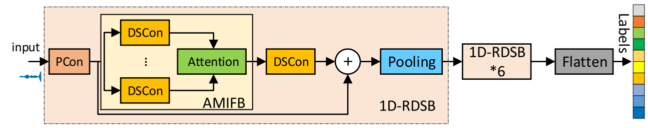

Fig. 1 Overall architecture of the proposed EMRNet. Here, PCon denotes pointwise convolution, DSCon represents depth separable convolution, and AMIFB is the adaptive multi-scale information fusion block, which will be detailed described in Fig. 2.

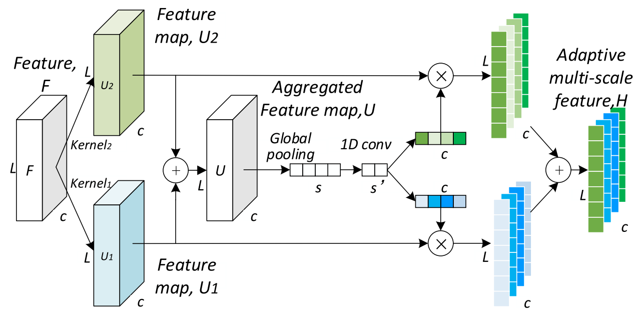

Fig. 2 Flowchart of the proposed adaptive multi-scale information fusion block (AMIFB), where L and c denote the length and thickness of the feature map, respectively.

3.1 1D Depthwise Separable Convolutions

Inspired by the MobileNet [15], the depth separable convolution technique is applied in the proposed 1D-RDSB. To build a more efficient neural network architecture, the depth separable convolution technique replaces standard convolution with a combination of depthwise convolution and pointwise convolution. Depthwise convolution utilizes a single-layer filter to each input channel for spatial filtering. Then, pointwise convolution maps features through computing a linear aggregation of spatial filtering features.

3.2 Adaptive Multi-Scale Information Fusion

CNN is inspired by local receptive fields of neurons. In traditional CNN structures, the size of local receptive fields is fixed, which results in the limited scale spatial information of the extracted features. To collect information with different scales, inception modules [14] concatenate a multi-branch structure, and each branch has a different size of receptive fields. However, these linear aggregation approaches ignore the relationship between different scale information. In the human cognitive system, multi-scale information from different sizes of receptive fields can be aggregated adaptively based on the importance of information. This behavior of concentrating the allocation of resources towards the most informative components is the attention mechanism [16], which has demonstrated its validness in adaptively aggregating channel information [17], [18]. Therefore, the adaptive aggregation of multi-scale features through the channels attention mechanism will dramatically enhance network learning ability [19]. As demonstrated in Fig. 2, adaptive channel attention is applied to fuse multi-scale information from multiple branches with different receptive fields.

3.3 Deep Residual Learning

Deep residual learning and skip connection are further embedded in the proposed network. [8], [20] has confirmed that residual learning can solve the problem of gradient disappearance existing in the training process and meanwhile strengthen network learning capability.

4. Results and Analysis

In this section, experiments based on ten kinds of CW radar signals are conducted to verify the performance of the designed EMRNet. The evaluation indicators adopted in our experiments include recognition accuracy, the number of multiply-accumulate operations (MACC), and paraments of the network. LFM, SFM, BPSK, Frank, FSK 4FSK, BASK, NS, LFM-BPSK, and SFM-BPSK are considered in experiments. The length of the received signal is set as $L = 1024$, which is the same as [5], [7], [20] for a fair comparison. The sampling rate utilized in the experiment is 400 MHz. Signal parameters are set with the normalized sampling frequency and the normalized length of the signal to demonstrate the generalizability of the model. The carrier frequency of all signals is distributed in $(0.1 \sim 0.4) * f_s$. The range of bandwidth of FM signals is in $(0.1 \sim 0.2) * f_s$. To guarantee that the intercepted signal contains at least one complete time interval, the number of samples per signal period is set as $(1/3 \sim 1/2) * L$. For BPSK, the Barker codes length is any of 5, 7, 11, and 13. BASK, FSK, and 4FSK adopt the random code. The phase number of polyphase code signals is from 4 to 7. From -4 dB to 20 dB, each signal generates 500 samples at a step of 2 dB for model training. In the testing set, 100 samples per 2 dB for each signal are produced from -10 dB to 10 dB. There are a total of 65000 samples in the training set and 11000 samples in the testing set. The optimization utilized in experiments is Adam, and the batch size is 128. The learning rate is 0.001 at the first 100 epochs, and then the learning rate is reduced tenfold every 50 epochs. A total of 200 epochs have been run.

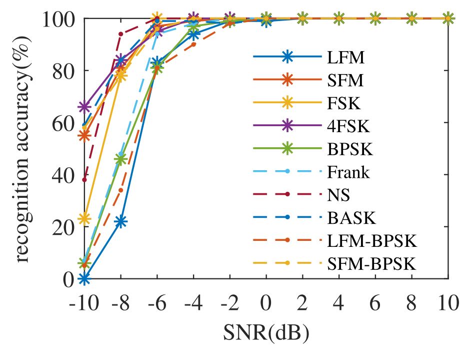

Fig. 3 Recognition accuracy of ten kinds of signals.

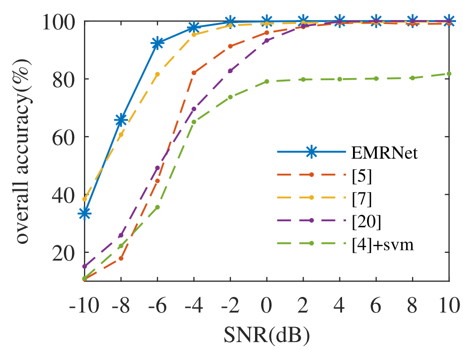

Fig. 4 Comparison with other methods.

Table 1 Complexity and size of different models

Method	MACC	Parameters
LeNet-5 [5]	1,890,792	114,106
SResNet-50 [7]	23,289,664	1,497,514
1D-ResNet [20]	42,419,456	212,490
EMRNet	1,671,680	18,516

The first experiment shows the recognition accuracy of the proposed EMRNet for ten types of CW radar signals in detail. Figure 3 indicates that the recognition accuracy of various signals is positively correlated with SNR. All signals maintain a 100% recognition performance when SNR exceeds 0 dB. The precision of LFM, LFM-BPSK, BPSK, and Frank signals declines sharply once the SNR is below 0 dB. While EMRNet can identify SFM, NS, FSK, 4FSK, BASK, and SFM-BPSK signals with a precision of more than 95% at the SNR of -6 dB. EMRNet collects multi-scale information with different receptive fields, providing a stronger multi-scale representation ability. Moreover, the adaptive multi-scale information aggregation effectively compensates for the loss of single-scale information caused by intense noise, which significantly improves the robustness of the model under low SNR conditions.

In order to exhibit the excellent performance of the designed approach, comparative experiments with other state-of-the-art algorithms are shown in Fig. 4. In [4], handcrafted features are extracted from the ambiguity function to classify radar signals. LeNet-5, consisting of two convolutional layers, is proposed in [5] to identify the TFIs of radar signals. In [7], shrinking ResNet-50 (SResNet) is adopted. In addition, the 1D-ResNet [20] based on I/Q samples is added to the comparison. For the traditional machine learning (ML) method [4], recognition precision is low due to the limitations of handcraft features. The accuracy of 1D-ResNet [20] is significantly higher than that of the ML method. Nevertheless, due to the poor robustness of the IQ samples, the classification effect is reduced considerably when the SNR is lower than 0 dB. Deep learning methods [5], [7] based on 2D CNN can recognize various radar signals with relatively high accuracy, but it is still not as good as the proposed method. In addition, the EMRNet has 1,671,680 MACCs, which is reduced by 13.93× compared with SResNet. 1D-ResNet also holds a 25.37× amount of computation than EMRNet. The fewer MACCs mean a lower network delay. This allows EMRNet to complete tasks efficiently, even on computationally limited platforms. Last but not least, EMRNet has 18,516 parameters, which is much less than other models. Such a tiny model of EMRNet can be embedded in almost any device. Comparative experiments reveal that EMRNet has both accuracy and efficiency, which is significantly better than previous methods.

5. Conclusion

In this letter, we proposed an efficient model, namely EMRNet, for CW radar waveform recognition. In the designed EMRNet, 1D depthwise separable convolutions and adaptive multi-scale information feature fusion techniques are first integrated to efficiently trade-off latency and accuracy. Comparative experiments verify that EMRNet can realize better recognition accuracy than other state-of-the-art methods while maintaining an extraordinarily computation-efficient and lightweight architecture.

References

[1] S. Liu, X. Yan, P. Li, X. Hao, and K. Wang, "Radar emitter recognition based on SIFT position and scale features," IEEE Trans. Circuits Syst. II, Exp. Briefs, vol.65, no.12, pp.2062–2066, Dec. 2018.

[2] F. Uysal, "Phase-coded FMCW automotive radar: System design and interference mitigation," IEEE Trans. Veh. Technol., vol.69, no.1, pp.270–281, Jan. 2020.

[3] J. Lunden and V. Koivunen, "Automatic radar waveform recognition," IEEE J. Sel. Topics Signal Process., vol.1, no.1, pp.124–136, June 2007.

[4] C. Xu, J. Zhang, Q. Zhou, S. Chen, "Recognition of radar signals based on AF grids and geometric shape constraint," Signal Process., vol.157, pp.30–34, April 2019.

[5] M. Zhang, M. Diao, and L. Guo, "Convolutional neural networks for automatic cognitive radio waveform recognition," IEEE Access, vol.5, pp.11074–11082, 2017.

[6] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proc. IEEE, vol.86, no.11, pp.2278–2324, Nov. 1998.

[7] K. Chen, S. Zhang, L. Zhu, S. Chen, and H. Zhao, "Modulation recognition of radar signals based on adaptive singular value reconstruction and deep residual learning," Sensors, vol.21, no.2, Jan. 2021.

[8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," Proc. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.770–778, IEEE, 2016.

[9] Z. Zhang, C. Wang, C. Gan, S. Sun, and M. Wang, "Automatic modulation classification using convolutional neural network with features fusion of SPWVD and BJD," IEEE Trans. Signal Inf. Process. Networks, vol.5, no.3, pp.469–478, Sept. 2019.

[10] K. Chen, L. Zhu, S. Chen, S. Zhang, and H. Zhao, "Deep residual learning in modulation recognition of radar signals using higher-order spectral distribution," Measurement, vol.185, Article No. 109945, Nov. 2021.

[11] Z. Xiao and Z. Yan, "Radar emitter identification based on novel time-frequency spectrum and convolutional neural network," IEEE Commun. Lett., vol.25, no.8, pp.2634–2638, Aug. 2021.

[12] T. Huynh-The, V.-S. Doan, C.-H. Hua, Q.-V. Pham, T.-V. Nguyen, and D.-S. Kim, "Accurate LPI radar waveform recognition With CWD-TFA for deep convolutional network," IEEE Wireless Commun. Lett., vol.10, no.8, pp.1638–1642, Aug. 2021.

[13] H. Chen and W. Ye. "Classification of human activity based on radar signal using 1-D convolutional neural network," IEEE Geos. Remote Sens. Lett., vol.17, no.7, pp.1178–1182, July 2020.

[14] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp.1–9, June 2015.

[15] A.G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," https://arxiv.org/abs/1704.04861, April 2017.

[16] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," Comput. Sci., 2014.

[17] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, "Squeeze-and-Excitation Networks," IEEE Trans. Pattern Anal. Mach. Intell., vol.42, no.8, pp.2011–2023, Aug. 2019.

[18] S. Woo, J. Park, J. Lee, and I. Kweon, "CBAM: Convolutional block attention module," ECCV, Oct. 2018.

[19] X. Li, W. Wang, X. Hu, and J. Yang, "Selective Kernel Networks," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp.510–519, 2019.

[20] 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.