IEEE Access
Multidisciplinary ⁝ Rapid Review ⁝ Open Access Journal

Received September 22, 2020, accepted October 19, 2020, date of publication October 27, 2020,
date of current version November 11, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3033989

Lightweight Deep Learning Model for Automatic Modulation Classification in Cognitive Radio Networks

SEUNG-HWAN KIM, JAE-WOO KIM, VAN-SANG DOAN, AND DONG-SEONG KIM, (Senior Member, IEEE)
Department of ICT Convergence Engineering, Kumoh National Institute of Technology, Gumi 39177, South Korea

Corresponding author: Dong-Seong Kim (dskim@kumoh.ac.kr)

This work was supported in part by the MSIT (Ministry of Science, ICT), Korea, under the Grand Information Technology Research Center Support Program, supervised by the IITP (Institute for Information and Communications Technology Planning and Evaluation), under Grant IITP-2020-2020-0-01612, in part by the Priority Research Centers Program under Grant 2018R1A6A1A03024003, and in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology, under Grant 2019R1I1A1A01063895.

ABSTRACT Automatic modulation classification (AMC) used in cognitive radio networks is an important class of methods apt to utilize spectrum resources efficiently. However, conventional likelihood-based approaches have high computational complexity. Thus, this paper proposes a novel convolutional neural network architecture for AMC. A bottleneck and asymmetric convolution structure are employed in the proposed model, which can reduce the computational complexity. The skip connection technique is used to solve the vanishing gradient problem and improve the classification accuracy. The dataset DeepSig:RadioML, which is composed of 24 modulation classes, is used for the performance analysis. Simulation results show that the classification accuracy performance of the proposed model is outstanding in the signal-to-noise ratio (SNR) range from -4 *dB* to 20 *dB* compared with MCNet that is the best model in the conventional models, where the proposed model achieves 5.52% and 5.92% improvement regarding classification accuracy at the SNRs of 0 *dB* and 10 *dB*, respectively. In terms of the computational complexity, the proposed model not only saves the trainable parameters by more than 67% but also reduces the prediction time for a signal by more than 54.4% compared with those of MCNet.

INDEX TERMS Automatic modulation classification, deep learning model, convolution neural network, light weight, cognitive radio.

I. INTRODUCTION

Automatic modulation classification (AMC) is widely used in military and industrial applications. With the rapid development of wireless communication technologies and the increasing demand for wireless services [1], the spectrum resources for wireless communications are becoming rapidly exhausted [2], [3]. An efficient approach to deal with this problem is to use cognitive radio (CR) technology, which significantly improves the spectrum utilization efficiency by sharing the licensed band between licensed and unlicensed users [4], [5]. In CR networks, the AMC method should be employed to identify a received signal without prior knowledge of the signal in various unknown channels. The AMC methods are generally classified into likelihood-based (LB) and feature-based (FB) approaches [6]. The former method maximizes the classification accuracy. However, it has a high computational complexity; consequently, its implementation in real-time applications is difficult [7]. The latter method consists of two stages: feature extraction and classification. One of the tasks in the feature extraction stage is the estimation of key factors such as the carrier frequency and signal power [8]. In the classification stage, higher-order statistics (HOS) is generally used as a classifier, which makes a decision based on statistical feature values [9], [10]. In addition, machine learning techniques such as support vector machine and decision tree can be applied [11], [12]. Consequently, the FB approach has relatively higher robustness than

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

the LB approach because of its high efficiency in implementation, although it yields a suboptimal performance. However, the FB approach has limitations in extracting expert knowledge from multiple signals because of high-order modulation types and impairment channel conditions. Moreover, the selection of feature values depends on a manual analysis [13].

Recently, deep learning techniques have attracted attention for their promising performance in diverse applications [14]–[16], which is a branch of machine learning [17]. A convolutional neural network (CNN) is a state-of-the-art, effective deep learning technique. Typical CNN models include AlexNet [18], GoogleNet [19], VGG [20], and ResNet [21]. However, these models focus on image processing [22], [23] and computer vision [24], [25] applications to classify images, and consequently, they are computationally heavy because of deep layers and large filter sizes. Moveover, they take a long time to predict each class. Therefore, the traditional architecture needs to be redesigned for AMC because real-time classification should be considered to satisfy the quality of service (QoS) requirement for CR networks [26]–[28]. For example, the time slot periods of the IEEE 802.15.4 and IEEE 802.11n standards are at least a few milliseconds and microseconds, respectively. Therefore, in addition to the conventional propagation, queuing, and processing delays, the classification delay of AMC influences the QoS. Thus, a CNN-based model may be promising for AMC because of its high classification capability [29]–[31].

Several CNN-based architectures have been proposed with the development of deep learning techniques [32]–[40]. For example, J. Shi *et al.* [32] proposed a deep-learning-based AMC model, which considered the phase offset effect for a realistic orthogonal frequency-division multiplexing (OFDM) system. In an OFDM system, four modulation schemes were used and the fast Fourier transform (FFT) was applied to convert the data between the frequency and time domains. The proposed model achieved a high classification accuracy as compared with three other schemes. However, only four classes were used in the classification. The number of parameters used in the proposed networks, by using the 256-point inverse FFT, is over $2M$, which is too large. S. Hong *et al.* [33] presented a CNN-based model for an OFDM system that is almost similar to [32]. The proposed network showed a good performance except at a low signal-to-noise ratio (SNR). However, although the authors used a small in-phase and quadrature (IQ) length and focused on only four classes, the architecture was computationally too heavy. H. Zhang *et al.* [34] proposed a deep multistream neural network for AMC. The proposed network was similar to GoogleNet and deployed several modules to enhance the classification accuracy. Furthermore, the authors used a 1 × 1 convolution layer to reduce the computational complexity of the networks. However, although the accuracy performance was good, more than $528k$ parameters were used in the proposed structure, which is still large relative to the small IQ length. S. Zheng *et al.* [13] proposed a fusion-based CNN architecture for AMC. The authors also presented two basic models (CNN1 and CNN2) for these fusion models, which use 2 and 46 convolutional layers, respectively. Both the basic and fusion models for 12 classes showed good classification performances. However, the CNN1 model used a very large number of parameters, i.e., over $21M$, and the CNN2 model used over $5M$ parameters. Therefore, when N numbers of each model are used for fusion, the number of parameters increases by N times. H. Gu *et al.* [35] proposed two types of CNN-based models for AMC in general scenarios. The first one was for blind channel identification, which predicted line-of-sight and non-line-of-sight classes, whereas the second one was for modulation classification, which predicted six modulation classes. These models used very few parameters, i.e., below $110k$, and showed good classification performance. However, the first model was used to classify very simple patterns, and the second model was used to classify low-order modulation schemes (2FSK, DQPSK, 16QAM, 4PAM, MSK, and GMSK), which are easy to recognize. C. Yang *et al.* [36] proposed a deep-learning-aided architecture for AMC. The proposed model was the same as the second model [35] mentioned earlier, and the target modulations for classification were also the same. Therefore, this proposed model has the same limitations as the second model mentioned above [35]. Y. Wang *et al.* [37] proposed two different deep-learning-based models for AMC. One was trained using IQ component signals, whereas the other was trained using image-based constellation diagrams. The IQ-based model had a computationally lighter architecture than the image-based model, and it showed a good performance at low- and high-SNR. However, it predicted a relatively small number of classes and the number of parameters used in the model was not small when the IQ length was considered. The image-based model could not be compared with the IQ-based model because its input size was larger. However, it had deeper layers, and used a larger number of parameters, even though it showed a better performance.

TABLE 1. **Summarization of related works.**

Ref.	Used param.	Classes	Frame length
J. Shi *et al.* [32]	$2M$	4	256
S. Hong *et al.* [33]	$2M$	4	256
H. Zhang *et al.* [34]	$528k$	10	256
S. Zheng *et al.* [13]	$21M$	12	512
H. Gu *et al.* [35]	$110k$	6	256
C. Yang *et al.* [36]	$110k$	6	256
Y. Wang *et al.* [37]	$1.1M$	7	128

In this context, we propose a lightweight model using a deep-learning-based CNN for CR networks. The DeepSig: RadioML 2018.01A dataset [39], which includes 24 modulation classes, is used for performance analysis. The following are the contributions of this paper:

1) The proposed architecture is designed to be computationally light by applying a bottleneck and asymmetric

TABLE 2. Key notations.

Symbol	Description
h	Rayleigh fading channel coefficient
x	Modulated signal
n	Additive white Gaussian noise
α	Channel impulse response
Δf	Carrier frequency offset
$\Delta\phi$	Phase offset
Q	Width length of input size
P	Height length of input size
C	Input channel number
K	Input size
W	Convolutional kernel size
L_q	Width length of feature map size
L_p	Height length of feature map size
μ_b	Average of mini batch size
σ_b^2	Variance of mini batch size
δ	Numerical stability coefficient
γ	Scale factor
β	Shift factor
g_κ	Feature map κ element
g_χ	Pre-activation output

convolution structure. Therefore, it uses a smaller number of trainable parameters compared with the conventional architectures, which can reduce the computational complexity

2) The classification accuracy performance of the proposed model is improved by using the skip connection approach. The proposed model outperforms the conventional models in the SNR range from -4 dB to 20 dB.
3) A dataset having 24 classes is applied to demonstrate the robustness of the proposed model. In this dataset, as the number of classes increases, the classification level becomes higher. Finally, the proposed model shows outstanding performance in terms of the classification accuracy and computation complexity.

The rest of this paper is organized as follows: Section 2 describes the system model, Section 3 represents our proposed CNN architecture, Section 4 presents the simulation results and performance analysis, and Section 5 concludes the paper.

II. SYSTEM MODEL

A. SIGNAL MODEL

In this study, DeepSig:RadioML 2018.01A datasets generated using USRP B210 are used (SNR from -10 dB to $+20$ dB), where the Rayleigh fading channel model is applied and other wireless communication conditions regarding random variables such as roll-off factor, carrier frequency offset, and symbol rate offset between the transceiver and the receiver can be found in [39] in details. The transmitted signals are sent as a frame, which is composed of 1024 samples, and the transceiver uses the 900MHz industrial, scientific, and medical (ISM) band. The system model, which is a 2 × 2 multiple-input multiple-output (MIMO) system for example, is shown in Fig. 1. The kth received signals in general form can be represented as

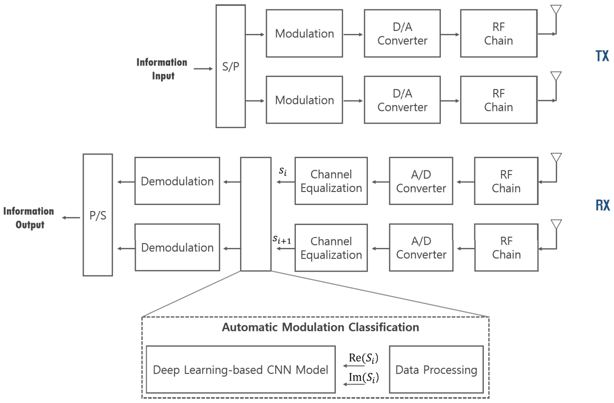


FIGURE 1. Example of system model.

$$\begin{aligned} y_1 &= h_{11}x_{11} + h_{21}x_{21} + n, \\ y_2 &= h_{22}x_{22} + h_{12}x_{12} + n, \end{aligned} \tag{1}$$

where the first index of h and x is the transmitted antenna number, the second index is the received antenna number, h is the Rayleigh fading channel coefficient, x is the modulated signal, and n is additive white Gaussian noise. In addition, h can be expressed as

$$h = \alpha e^{j(2\pi\Delta ft + \Delta\phi)}, \tag{2}$$

where α is the channel impulse response, and Δf and $\Delta\phi$ are the carrier frequency and phase offset, respectively. The offsets are mainly due to disparate local oscillators and the resultant Doppler effect in wireless communication. The channel loss from the received signals is compensated by channel equalization (CE), where the final input from CE is represented by a complex envelop. To utilize the received signal as the input value of the CNN model, it should be normalized by using root mean square (RMS), which prevents the optimizer from halting at the local optimal value, and then divided into IQ components. This is undertaken in the data processing stage.

B. DATA PROCESSING

Data processing consists of two tasks: normalization and high-dimensional representation for an IQ size of 2 × 1024. The datasets have 24 modulation types. Each modulation has 4096 frames, which are used for training (80%) and testing (20%), and the modulations are divided into two groups, namely, digital and analog modulations, as follows:

1) Digital modulation: OOK, 4ASK, 8ASK, BPSK, QPSK, OQPSK, 8PSK, 16PSK, 32PSK, 16APSK, 32APSK, 64APSK, 128APSK, 16QAM, 32QAM, 64QAM, 128QAM, 256QAM, GMSK

2) Analog modulation: FM, AM-SSB-WC, AM-SSB-SC, AM-DSB-WC, AM-DSB-SC

The RMS scheme is applied to the complex envelop signal for normalization, which is represented as

$$\bar{s}_i = \frac{s_i}{RMS(s)}, \quad where\ RMS(s) = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1}|s_i|^2}, \tag{3}$$

where N is the number of samples in a frame, which is 1024 in total. After normalization, the samples are aggregated into a $2 \times N$ matrix, which is represented as

$$S_j = \begin{bmatrix} \bar{s}_0^I & \bar{s}_1^I & \cdots & \bar{s}_{N-1}^I \\ \bar{s}_0^Q & \bar{s}_1^Q & \cdots & \bar{s}_{N-1}^Q \end{bmatrix}, \tag{4}$$

where superscript I is the in-phase term of $\bar{s}_i$, which is equal to $Re[\bar{s}_i]$, and superscript Q is the quadrature term of $\bar{s}_i$, which is equal to $Im[\bar{s}_i]$. Finally, according to the number of frames and modulation types, the data matrix based on S_j for the CNN is represented as

$$H = \begin{bmatrix} \bar{S}_{1,1} & \bar{S}_{1,2} & \cdots & \bar{S}_{1,k} \\ \bar{S}_{2,1} & \bar{S}_{2,2} & \cdots & \bar{S}_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{S}_{m,1} & \bar{S}_{m,2} & \cdots & \bar{S}_{m,k} \end{bmatrix}, \tag{5}$$

where subscript m is the total number of frames and subscript k is the total number of modulations.

III. PROPOSED CNN ARCHITECTURE

A. CNN ARCHITECTURE

In general, a CNN architecture is composed of a convolutional, batch normalization (BN), activation function, pooling, fully connected, and classification layers. The convolutional layer is used to extract feature maps. For example, when the input size is $Q \times P \times C$, where C is the input channel number, the feature map can be represented [40] as

$$z^l = \sum_{c=1}^{C}\sum_{\rho=0}^{\Gamma-1}\sum_{\upsilon=0}^{\Upsilon-1} K(q+\rho, p+\upsilon, c) \times W^l(\rho, \upsilon) + b^l, \tag{6}$$

where $W^l(.)$ is the l_{th} convolutional kernel size, b^l is the l_{th} bias, and K is the input size, where q and p denote the row and column index of the input size, respectively. In addition, $q = [1, L_q]$ where $L_q = [\frac{Q-\Gamma}{stride} + 1]$, and $p = [1, L_p]$ where $L_p = [\frac{P-\Upsilon}{stride} + 1]$. The feature map passed from the convolutional layer is normalized by the BN layer, which is represented as

$$\acute{z^l} = \frac{z^l - \mu_b}{\sqrt{\sigma_b^2 + \delta}}, \tag{7}$$

$$g^l = \gamma \acute{z^l} + \beta, \tag{8}$$

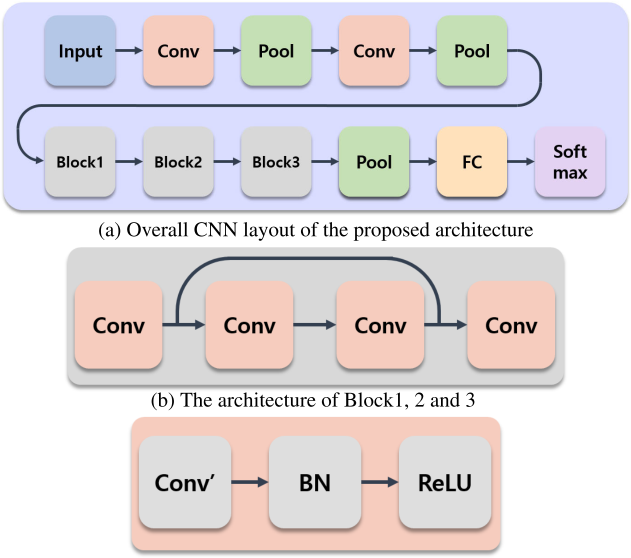
where μ_b and σ_b^2 are the average and variance of the mini-batch size, respectively. δ is a numerical stability coefficient, which prevents the denominator from becoming zero. γ is a scale factor and β is a shift factor over the mini-batch size. The rectified linear unit (ReLU) activation function decides whether the output of the BN is forwarded to the next layer, and it is expressed as

$$f_{ReLU(\acute{g_l})} = max(0, g_l), \tag{9}$$

where the softmax activation function is used after the fully connected layer, which is given as

$$f_{softmax(g_\kappa^l)} = \frac{e^{g_\kappa^l}}{\sum_\chi e^{g_\chi^l}}, \tag{10}$$

where g_κ^l is the l_{th} feature map κ element and g_χ^l is a pre-activation output. Finally, stochastic gradient descent with momentum (SGDM) [41] is applied as the cost function for gradient descent optimization in this study.



(a) Overall CNN layout of the proposed architecture

(b) The architecture of Block1, 2 and 3

(c) The structure of Conv block

FIGURE 2. Description of the proposed network architecture.

B. PROPOSED CNN ARCHITECTURE

As shown in Fig. 2, the lightweight CNN architecture is composed of 10 convolutional layers, 3 pooling layers, and 1 fully connected layer, where *Conv* and *Pool* represent the convolution and average pooling operations, respectively. In addition, each convolution block is composed of a convolutional layer, BN, and ReLU activation function in Fig. 2c. The first convolutional layer processes the input dimension 2 × 1024 from the input layer, where the convolutional layer contains a 3 × 3 kernel matrix, and 16 convolution kernels, and the padding size (3, 1) is applied. The first pooling layer reduces the size of the feature map to optimize the extraction of the signal characteristics, where the kernel size is 2 × 2. The convolution and pooling operations are again performed with 32 convolution kernels, where the size of the feature map is optimized to obtain sufficient signal characteristics for a high performance accuracy. Subsequently, the output feature map passes Blocks 1, 2, and 3 in succession without downsampling, where the first and last convolutional layers of Blocks 1, 2, and 3 use a 1 × 1 kernel matrix to reduce the dimension of the feature map channel, which affects the computational complexity.

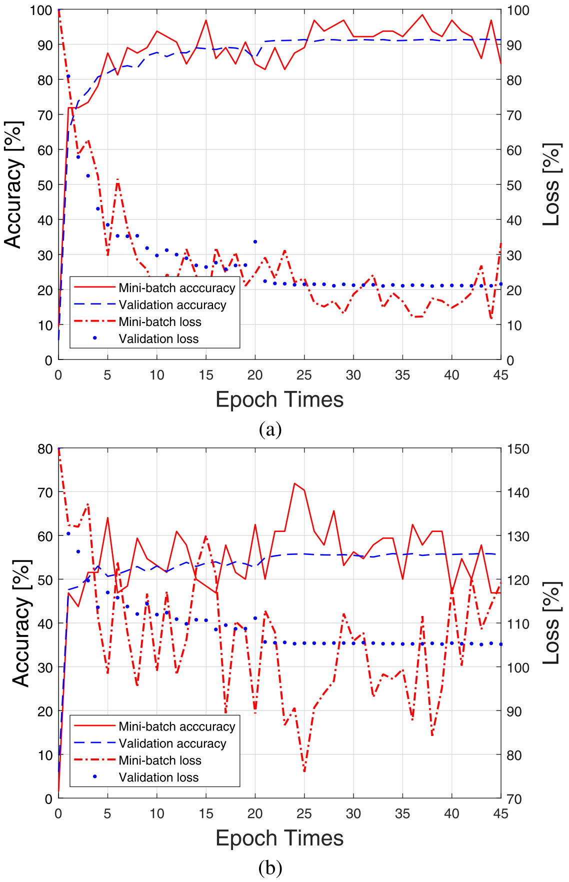
Furthermore, the second and third convolutional layers use an asymmetric kernel matrix of kernel sizes 3 × 1 and 1 × 3, respectively, instead of 3 × 3 to decrease the number of trainable parameters. Therefore, when the proposed kernels are applied instead of a kernel size of 3 × 3 in Block 1, 12*k* trainable parameters can be saved. Moreover, 32k trainable parameters can be saved in the case of Block 2, whereas 129*k* trainable parameters can be saved in the case of Block 3. The accuracy performance can be enhanced by using the skip connection technique for the vanishing gradient problem. It is applied to Blocks 1, 2, and 3 as shown in Fig. 2b. The last pooling layer reduces the size of the feature map from 3 × 256 to 1 × 1 to prevent overfitting before the classification step, which includes the fully connected and softmax layers. The final feature map is connected by a fully connected layer, whose output is the same as the number of modulations. The detailed information of the proposed model is shown in Table 3, which is including the output size, the filter size, and the number of used parameters. As shown in Table 3, the parameters are 46472, which are used in the model.

TABLE 3. Configuration of the proposed CNN architecture.

Type	Output Size	Filter Size	Parameters
Input	2 × 1024	-	-
Conv	6 × 1024 × 16	16 × 3 × 3	192
Pool	2 × 512 × 16	2 × 2	-
Conv	6 × 512 × 32	32 × 3 × 3	4704
Pool	3 × 256 × 32	2 × 2	-
Block1	3 × 256 × 32	32 × 1 × 1	1120
	3 × 256 × 8	8 × 3 × 1	792
	3 × 256 × 8	8 × 1 × 3	216
	3 × 256 × 32	32 × 1 × 1	352
Block2	3 × 256 × 64	64 × 1 × 1	2240
	3 × 256 × 16	16 × 3 × 1	3120
	3 × 256 × 16	16 × 1 × 3	816
	3 × 256 × 64	64 × 1 × 1	1216
Block3	3 × 256 × 128	128 × 1 × 1	8576
	3 × 256 × 32	32 × 3 × 1	12384
	3 × 256 × 32	32 × 1 × 3	3168
	3 × 256 × 128	128 × 1 × 1	4480
Pool	1 × 1 × 128	3 × 256	-
FC	24 × 128	-	3096
Softmax	24	-	-

IV. SIMULATION RESULTS

For simulation works, the SNR range of the used dataset is from −10 *dB* to 20 *dB* with the interval of 2 *dB*. Therefore, for the simulation work of each SNR domain 98304 frames are used, which is composed of one of 24 modulations (4096 frames). Finally, the dataset gets 1572864 frames, where the dataset is divided such that 80% of the frames (1258291 frames) are used for training and the remaining (314573 frames) are used for testing. The performance of the proposed model is compared with that of conventional models such as ML-XGboost [39], VGG [39], ResNet [39], CNN-AMC [42], and MCNet [43]. The simulation is performed using an i5 2.9 GHz CPU, 32 GB RAM, and NVIDIA GeForce RTX 2080 Super GPU devices. The detailed configuration of the simulation is summarized in Table 4.



(a)

(b)

FIGURE 3. Mini-batch/validation accuracy and loss versus epoch times. (a)SNR=10 dB;(b)SNR=0 dB.

TABLE 4. Configuration for the simulation.

Type	Value
MaxEpochs	45
MiniBatchSize	64
InitialLearningRate	0.01
LearningRateDropPeriod	20
LearningRateDropFactor	0.01
Optimizer	SGDM

The mini-batch/validation accuracy and loss of the proposed model is represented (Fig. 3). As the results, the convergence of validation accuracy and loss is reached roughly from 21 epoch times, and it is shown that they are converged stably. In terms of the mini-batch accuracy and loss in Fig. 3b, the curves are fluctuated unstably because the initial learning rate is set relatively high. For learning of the proposed model it takes average 65 minutes on NVIDIA RTX 2080 Super. When compared with Oshea *et al.* [39], the lead time is longer due to the difference of GPU performance.

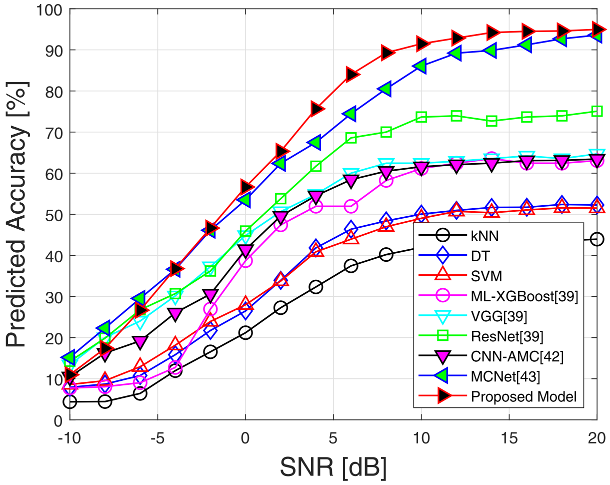


FIGURE 4. **The performance comparison with the conventional models and the proposed model.**

The classification accuracy for each model is represented (Fig. 4), and the conventional models are compared with the proposed model. From the results, the proposed model outperforms the conventional models within the SNR range of −4 and 20 *dB*. However, between −6 and −10 *dB*, MCNet performs better than the proposed models and exhibits the best performance among the conventional models. As classical methods, kNN (k-Nearest Neighbor), DT (Decision Tree), and SVM (Support vector machine) are used. To apply a feature set to the classifiers there are instantaneous features [44], HOC (High-Order Cumulant) features [10], and cyclostationary features [45], [46], where instantaneous features are used for a feature set in the classical methods. As the result, the classical methods are no better than deep leaning-based classifiers, and kNN is the worst in the classical methods. The ML-XGBoost model, in which an advanced DT algorithm is used, still shows lower performance than deep learning-based model. In cases where the deep learning technique is applied, such as the MCNet and ResNet models, a significant difference in accuracy with the VGG and CNN-AMC models was observed because the skip connection technique is used in the VGG and CNN-AMC models. At an SNR of 10 *dB*, the proposed model achieves 5.92% and 33.17% improvement compared with the MCNet and ML-XGBoost models, respectively. At a 0 *dB* SNR, the proposed model is higher than the MCNet and ML-XGBoost models by 5.52% and 31.65%, respectively.

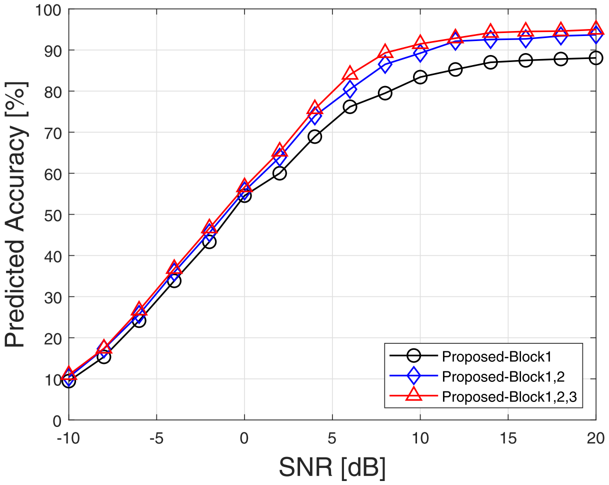
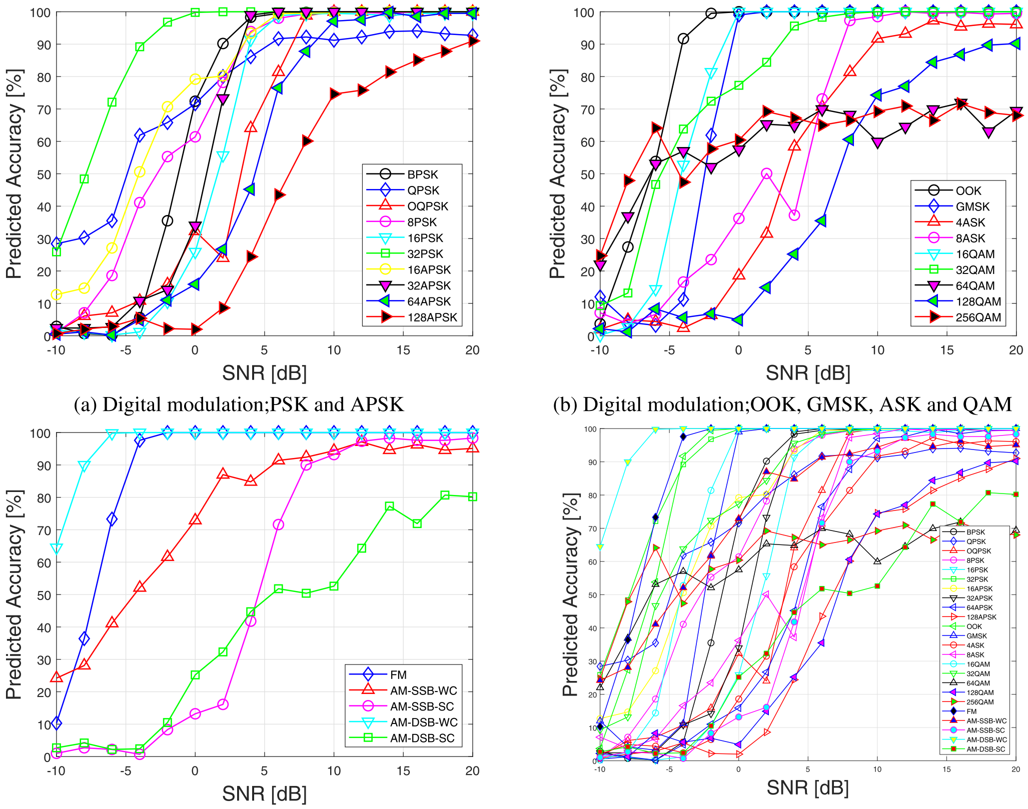


FIGURE 5. **Classification result in the different number of blocks applied to the proposed model.**

In Fig. 5, the accuracy performance is illustrated when different numbers of blocks are applied to the proposed model, which has three blocks in total. When only Block 1 is applied in the SNR range from −10 *dB* to 0 *dB*, there is a slight gap in the accuracy performance. However, in the SNR range from 2 *dB* to 20 *dB*, the gap is relatively large. However, when Block 1,2 or Block 1,2,3 are applied, the gap in the accuracy performance is small in the overall SNR range.

In Fig. 6, 24 modulation schemes are shown in 3 groups. In Fig. 6a, the performance of the PSK and APSK modulation schemes is illustrated. The modulation schemes are divided equally into two categories, namely, low performance and high performance, at SNR 0 *dB*. The schemes in the low-performance category—which are OQPSK, 16PSK, 32APSK, 64APSK, and 128 APSK—are predicted with an accuracy below 40%. The schemes in the high-performance category—which are the remaining schemes—are predicted with an accuracy over 60%. In contrast, all the modulation schemes are predicted with an accuracy over 90% at SNR 20 *dB*. In Fig. 6b, the performances of OOK, GMSK, ASK, and QAM modulation schemes are shown. All the modulation schemes except three—namely, OOK, GMSK, 16QAM, 32QAM, 64QAM, and 256QAM—have a classification accuracy of over 55%. Unfortunately, the accuracy performance of 64QAM and 256QAM remains at approximately 70% up to SNR 20 *dB*, although the remaining modulation schemes have an accuracy over 90% at SNR 20 *dB*. In Fig. 6c, the performance of analog modulation schemes is shown. Here, the AM-SSB-SC and AM-DSB-SC schemes show a low accuracy performance below 30% at SNR 0 *dB*. Furthermore, the AM-DSB-SC scheme is difficult to predict over the entire range of SNR, compared with the other modulation schemes.

In Fig. 7, the accuracy performance of each modulation scheme is shown at SNRs 0 *dB* and 10 *dB*. Here, the average accuracies at SNRs 0 *dB* and 10 *dB* are 56.64% and 91.48%, respectively. In Fig. 7a, as 128APSK and 128QAM are the most difficult to predict, they cannot be distinguished



(a) Digital modulation;PSK and APSK

(b) Digital modulation;OOK, GMSK, ASK and QAM

(c) Analog modulation; FM, AM-SSB and AM-DSB

(d) Overall 24 modulations (digital and anology)

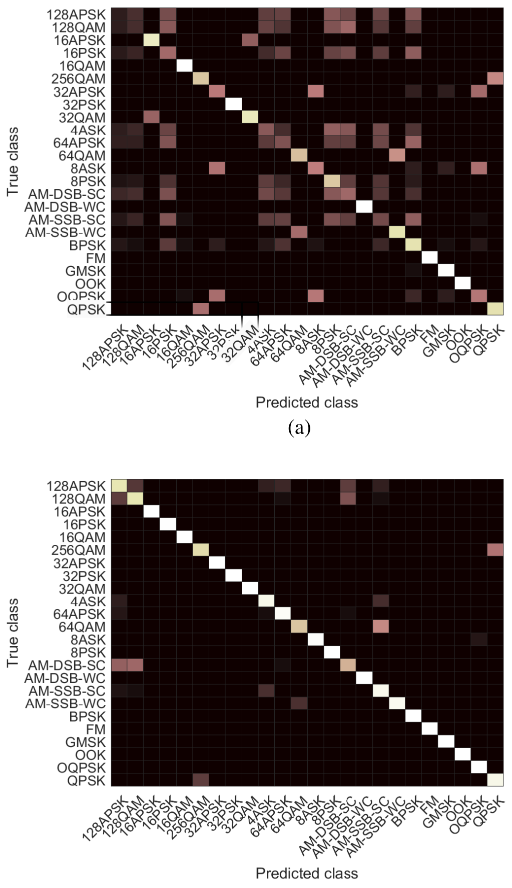
FIGURE 6. Classification results of digital and analog modulation.

TABLE 5. Summarization of classification accuracy.

SNR	kNN	DT	SVM	ML-XGBoost	VGG	ResNet	CNN-AMC	MCNet	Proposed
-10	4.39	7.90	8.60	7.74	13.39	14.23	10.47	15.19	10.97
-8	4.44	8.69	9.52	8.06	20.00	19.86	16.23	22.32	17.39
-6	6.42	10.78	12.90	9.03	24.03	26.76	19.19	29.53	26.66
-4	11.96	15.84	18.15	12.74	30.16	30.7	26.02	36.62	36.78
-2	16.54	21.73	23.92	26.94	37.26	36.2	30.59	46.12	46.65
0	21.19	26.49	28.00	38.71	44.84	45.92	41.43	53.51	56.64
2	27.27	33.93	33.76	47.42	50.65	53.80	49.58	62.38	65.32
4	32.32	41.82	40.87	51.94	54.84	61.69	54.51	67.48	75.65
6	37.45	46.37	43.94	51.94	60.00	68.59	58.42	74.45	84.01
8	40.20	48.39	46.97	58.23	62.42	70.00	60.49	80.52	89.29
10	41.95	50.04	49.02	61.13	62.42	73.66	61.57	86.06	91.48
12	42.46	50.95	50.72	62.42	62.9	73.94	62.07	89.20	92.86
14	42.27	51.66	50.38	63.55	63.55	72.68	62.44	89.86	94.21
16	43.46	51.71	51.05	62.42	64.19	73.66	63.06	91.23	94.5
18	43.52	52.39	51.56	62.42	63.55	73.94	63.11	92.69	94.59
20	43.94	52.26	51.50	63.06	64.68	75.07	63.44	93.59	94.97

from nine modulation schemes. In Fig. 7b, AM-DSB-SC and 64QAM are the most difficult to predict. Here, AM-DSB-SC cannot be distinguished from 128APSK and 128QAM, and 64QAM cannot be distinguished from AM-SSB-SC. The reference confusion matrices at 10 dB SNR can be referred by [43] To compare the proposed model with the baseline model (MCNet) regarding outstandingness. The overall classification accuracy for each model is summarized in Table 5.

The computational complexity of the models, which is an important factor in CR networks for real-time communication, is considered. The computational complexity is summarized in Table 6 for each model. In accordance with the



(a)

(b)

FIGURE 7. Confusion matrix of proposed model. (a) SNR=0 dB, Accuracy: 56.64%; (b) SNR=10 dB, Accuracy: 91.48%.

TABLE 6. Summarization of computational complexity.

Model	Num. parameters	Predicted time (ms)
kNN	-	0.0018
DT	-	0.0015
SVM	-	0.054
VGG [39]	$257k$	0.131
ResNet [39]	$236k$	0.146
CNN-AMC [42]	$575k$	0.127
MCNet [43]	$142k$	0.125
Proposed	$46k$	0.057

used trainable parameters CNN-AMC is the heaviest model due to a sharp rise in the fully-connected layer, which uses 55.3% and 58.9% more trainable parameters than VGG and ResNet, respectively. However, the prediction time is 3% and 13% shorter than VGG and ResNet because a few number of the convolutional layers are applied. VGG and ResNet are heavier than MCNet, which uses 44.7% and 39.8% more trainable parameters, respectively. However, the prediction time of them is similar with VGG and ResNet because both depth-wise concatenation operations and addition operations are performed many times by the structural character of MCNet. The proposed model is computationally the lightest model, which uses 67.6% fewer trainable parameters than MCNet because not only small and asymmetric kernels are applied, but the layer depth is not deep. In terms of the prediction time, the proposed model requires a relatively very short period compared with the deep learning-based models—which is 0.057 ms—for predicting a signal. Through the comparison of the prediction time, it is observed that the proposed model can save 54.4% more than MCNet, which shows the excellence of the proposed model. In contrast, ResNet requires the longest time period in the models, and it spends 60.9% more than the proposed model. A comparison with the classical models, the proposed model requires a longer period for prediction due to the simple structure of kNN and DT, which depends on the size of a dataset. The complexity of SVM is higher than kNN and DT because it should makes multiple classifiers according to the number of class of a dataset. In the case of training time for each model, it can be changed by the hardware performance. Further, if multiple tasks work on PC when the dataset is training the training time can be changed. Thus, the training time is not displayed.

V. CONCLUSION

In this paper, a new CNN architecture for AMC was proposed. The proposed architecture was designed by applying a bottleneck and asymmetric convolution structure, which can reduce the computational complexity, to consider the real-time communication for CR networks. The dataset DeepSig:RadioML 2018.01A, which has 24 modulation classes, was used for performance analysis. The simulation results showed that the classification accuracy performance of the proposed model was better in the SNR range from -4 dB to 20 dB, where the proposed model achieved classification accuracies of 5.52% and 5.92% at SNRs 0 dB and 10 dB, respectively. In terms of the computational complexity, the proposed model saved over 67% trainable parameters and reduced the prediction time for a signal by over 54.4% compared with those of MCNet. Finally, the robustness of the proposed model was demonstrated via a comparison with the conventional models. For future works we will propose a novel approach for the difficult modulation (e.g. M-QAM, M-APSK and M-PSK) to predict in the harsh environment based on this study.

REFERENCES

[1] P. T. A. Quang and D.-S. Kim, "Enhancing real-time delivery of gradient routing for industrial wireless sensor networks," *IEEE Trans. Ind. Informat.*, vol. 8, no. 1, pp. 61–68, Feb. 2012.

[2] I. Kakalou, K. E. Psannis, P. Krawiec, and R. Badea, "Cognitive radio network and network service chaining toward 5G: Challenges and requirements," *IEEE Commun. Mag.*, vol. 55, no. 11, pp. 145–151, Nov. 2017.

[3] S.-H. Kim, J.-W. Kim, and D.-S. Kim, "Energy consumption analysis of beamforming and cooperative schemes for aircraft wireless sensor networks," *Appl. Sci.*, vol. 10, no. 12, pp. 4374–4391, Jun. 2020.

[4] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[5] K. Won Choi, "Adaptive sensing technique to maximize spectrum utilization in cognitive radio," *IEEE Trans. Veh. Technol.*, vol. 59, no. 2, pp. 992–998, Feb. 2010.

[6] 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.

[7] F. Hameed, O. Dobre, and D. Popescu, "On the likelihood-based approach to modulation classification," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 5884–5892, Dec. 2009.

[8] C. Weber, T. Felhauer, and M. Peter, "Automatic modulation classification technique for radio monitoring," *Electron. Lett.*, vol. 51, no. 10, pp. 794–796, May 2015.

[9] M. Abdelbar, W. H. Tranter, and T. Bose, "Cooperative cumulants-based modulation classification in distributed networks," *IEEE Trans. Cognit. Commun. Netw.*, vol. 4, no. 3, pp. 446–461, Sep. 2018.

[10] A. Swami and B. M. Sadler, "Hierarchical digital modulation classification using cumulants," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 416–429, Mar. 2000.

[11] Y. Wang, J. Wang, W. Zhang, J. Yang, and G. Gui, "Deep learning-based cooperative automatic modulation classification method for MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4575–4579, Apr. 2020.

[12] S. Peng, H. Jiang, H. Wang, H. Alwageed, Y. Zhou, M. M. Sebdani, and Y.-D. Yao, "Modulation classification based on signal constellation diagrams and deep learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 3, pp. 718–727, Mar. 2019.

[13] 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.

[14] G. Aceto, D. Ciuonzo, A. Montieri, and A. Pescape, "Mobile encrypted traffic classification using deep learning: Experimental evaluation, lessons learned, and challenges," *IEEE Trans. Netw. Service Manage.*, vol. 16, no. 2, pp. 445–458, Jun. 2019.

[15] N. Samuel, T. Diskin, and A. Wiesel, "Learning to detect," *IEEE Trans. Signal Process.*, vol. 67, no. 10, pp. 2554–2564, May 2019.

[16] G. Aceto, D. Ciuonzo, A. Montieri, and A. Pescapé, "Toward effective mobile encrypted traffic classification through deep learning," *Neurocomputing*, vol. 409, pp. 306–315, Oct. 2020.

[17] B. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[18] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Int. Conf. Adv. Neural Infor. Process. Syst. (NIPS)*, 2013, pp. 1097–1105.

[19] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Boston, MA, USA, Jun. 2015, pp. 1–9.

[20] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *CoRR*, vol. abs/1409.1556, pp. 1–14, Sep. 2014.

[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.

[22] A. Maier, C. Syben, T. Lasser, and C. Riess, "A gentle introduction to deep learning in medical image processing," *Zeitschrift für Medizinische Physik*, vol. 29, no. 2, pp. 86–101, May 2019.

[23] L. Jiao and J. Zhao, "A survey on the new generation of deep learning in image processing," *IEEE Access*, vol. 7, pp. 172231–172263, 2019.

[24] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep learning for computer vision: A brief review," *Comput. Intell. Neurosci.*, vol. 2018, pp. 1–13, 2018, Art. no. 7068349.

[25] N. Akhtar and A. Mian, "Threat of adversarial attacks on deep learning in computer vision: A survey," *IEEE Access*, vol. 6, pp. 14410–14430, 2018.

[26] Z. Chen, N. Guo, and R. C. Qiu, "Demonstration of real-time spectrum sensing for cognitive radio," *IEEE Commun. Lett.*, vol. 14, no. 10, pp. 915–917, Oct. 2010.

[27] Z. Liang, S. Feng, D. Zhao, and X. S. Shen, "Delay performance analysis for supporting real-time traffic in a cognitive radio sensor network," *IEEE Trans. Wireless Commun.*, vol. 10, no. 1, pp. 325–335, Jan. 2011.

[28] A. Homayounzadeh and M. Mahdavi, "Quality of service provisioning for real-time traffic in cognitive radio networks," *IEEE Commun. Lett.*, vol. 19, no. 3, pp. 467–470, Mar. 2015.

[29] H. Wu, Y. Li, L. Zhou, and J. Meng, "Convolutional neural network and multi-feature fusion for automatic modulation classification," *Electron. Lett.*, vol. 55, no. 16, pp. 895–897, Aug. 2019.

[30] J. Nie, Y. Zhang, Z. He, S. Chen, S. Gong, and W. Zhang, "Deep hierarchical network for automatic modulation classification," *IEEE Access*, vol. 7, pp. 94604–94613, 2019.

[31] A. P. Hermawan, R. R. Ginanjar, D.-S. Kim, and J.-M. Lee, "CNN-based automatic modulation classification for beyond 5G communications," *IEEE Commun. Lett.*, vol. 24, no. 5, pp. 1038–1041, May 2020.

[32] J. Shi, S. Hong, C. Cai, Y. Wang, H. Huang, and G. Gui, "Deep learning-based automatic modulation recognition method in the presence of phase offset," *IEEE Access*, vol. 8, pp. 42841–42847, 2020.

[33] S. Hong, Y. Zhang, Y. Wang, H. Gu, G. Gui, and H. Sari, "Deep learning-based signal modulation identification in OFDM systems," *IEEE Access*, vol. 7, pp. 114631–114638, 2019.

[34] H. Zhang, Y. Wang, L. Xu, T. A. Gulliver, and C. Cao, "Automatic modulation classification using a deep multi-stream neural network," *IEEE Access*, vol. 8, pp. 43888–43897, 2020.

[35] H. Gu, Y. Wang, S. Hong, and G. Gui, "Blind channel identification aided generalized automatic modulation recognition based on deep learning," *IEEE Access*, vol. 7, pp. 110722–110729, 2019.

[36] C. Yang, Z. He, Y. Peng, Y. Wang, and J. Yang, "Deep learning aided method for automatic modulation recognition," *IEEE Access*, vol. 7, pp. 109063–109068, 2019.

[37] 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.

[38] S. Huang, L. Chai, Z. Li, D. Zhang, Y. Yao, Y. Zhang, and Z. Feng, "Automatic modulation classification using compressive convolutional neural network," *IEEE Access*, vol. 7, pp. 79636–79643, 2019.

[39] 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.

[40] Z. Qu, X. Mao, and Z. Deng, "Radar signal intra-pulse modulation recognition based on convolutional neural network," *IEEE Access*, vol. 6, pp. 43874–43884, 2018.

[41] N. Qian, "On the momentum term in gradient descent learning algorithms," *Neural Netw.*, vol. 12, no. 1, pp. 145–151, Jan. 1999.

[42] F. Meng, P. Chen, L. Wu, and X. Wang, "Automatic modulation classification: A deep learning enabled approach," *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, pp. 10760–10772, Nov. 2018.

[43] T. Huynh-The, C.-H. Hua, Q.-V. Pham, and D.-S. Kim, "MCNet: An efficient CNN architecture for robust automatic modulation classification," *IEEE Commun. Lett.*, vol. 24, no. 4, pp. 811–815, Apr. 2020.

[44] Z. Wu, S. Zhou, Z. Yin, B. Ma, and Z. Yang, "Robust automatic modulation classification under varying noise conditions," *IEEE Access*, vol. 5, pp. 19733–19741, 2017.

[45] O. A. Dobre, M. Oner, S. Rajan, and R. Inkol, "Cyclostationarity-based robust algorithms for QAM signal identification," *IEEE Commun. Lett.*, vol. 16, no. 1, pp. 12–15, Jan. 2012.

[46] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Cyclostationarity-based modulation classification of linear digital modulations in flat fading channels," *Wireless Pers. Commun.*, vol. 54, no. 4, pp. 699–717, Jul. 2009.

SEUNG-HWAN KIM received the B.E., M.S., and Ph.D. degrees in electronics engineering from the Kumoh National Institute of Technology (KIT), in 2010, 2012, and 2018, respectively. He currently works at the ICT Convergence Research Center, Kumoh National Institute of Technology, South Korea, as a Senior Researcher. His research interests are automatic modulation classification, deep learning, industrial wireless sensor networks, and the Internet of Things (IoT).

JAE-WOO KIM received the Ph.D. degree in computer engineering from the Kumoh National Institute of Technology (KIT). He worked for optical communications company ARTECH, from 2014 to 2018. He currently works at the ICT Convergence Research Center, Kumoh National Institute of Technology, South Korea, as a Research Professor. His research interests include optical transceiver, optical networks, the Internet of Things (IoT), and radio resource management in mobile communication systems.

VAN-SANG DOAN received the M.Sc. and Ph.D. degrees in electronic systems and devices from the Faculty of Military Technology, University of Defence, Brno, Czech Republic, in 2013 and 2016, respectively. He was three times awarded the Honors degrees by Faculty of Military Technology, University of Defence, in 2011, 2013, and 2016. He is currently a Postdoctoral Research Fellow with the ICT Convergence Research Center, Kumoh National Institute of Technology, South Korea. His current research interests include radar and sonar systems, signal processing, and deep learning.

DONG-SEONG KIM (Senior Member, IEEE) received the Ph.D. degree in electrical and computer engineering from Seoul National University, Seoul, South Korea, in 2003. From 1994 to 2003, he worked as a full-time Researcher at ERC-ACI, Seoul National University, Seoul. From March 2003 to February 2005, he worked as a Postdoctoral Researcher at the Wireless Network Laboratory, School of Electrical and Computer Engineering, Cornell University, NY. From 2007 to 2009, he was a Visiting Professor with the Department of Computer Science, University of California, Davis, CA. He is currently the Director of the KIT Convergence Research Institute and the ICT Convergence Research Center (Grand ICT program) supported by Korean government at the Kumoh National Institute of Technology. He is an ACM Senior Member. His current main research interests are real-time IoT, industrial wireless control networks, networked embedded systems, and fieldbus.