Deep Learning-Based Robust Automatic Modulation Classification for Cognitive Radio Networks

SEUNG-HWAN KIM [ID], **(Member, IEEE), JAE-WOO KIM,**
WILLIAMS-PAUL NWADIUGWU [ID]**, (Member, IEEE),**
AND DONG-SEONG KIM [ID]**, (Senior Member, IEEE)**
Department of IT Convergence Engineering, Kumoh National Institute of Technology, Gumi 39177, Republic of Korea

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

ABSTRACT In this paper, a novel deep learning-based robust automatic modulation classification (AMC) method is proposed for cognitive radio networks. Generally, as network input of AMC convolutional neural networks (CNNs) images or complex signals are utilized in time domain or frequency domain. In terms of the image that contains RGB(Red, Green, Blue) levels the input size may be larger than the complex signal, which represents the increase of computational complexity. In terms of the complex signal it is normally used as $2 \times N$ size for the input, which is divided into in-phase and quadrature-phase (IQ) components. In this paper, the input size is extended as $4 \times N$ size by copying IQ components and concatenating in reverse order to improve the classification accuracy. Since the increase in the amount of computation complexity due to the extended input size, the proposed CNN archiecture is designed to reduce the size from $4 \times N$ to $2 \times N$ by an average pooling layer, which can enhence the classification accuracy as well. The simulation results show that the classification accuracy of the proposed model is higher than the conventional models in the almost signal-to-noise ratio (SNR) range.

INDEX TERMS Automatic modulation classification, deep learning model, convolution neural network, frame extension, cognitive radio.

I. INTRODUCTION

Cognitive radio (CR) technology can make wireless devices connect one of unused spectrum subbands and exploit it [1], [2]. Thus, CR plays a important role to utilize the scarce spectrum resources, and satisfy the required spectrum demands due to advancements in internet-connected devices and wireless communication technology (e.g. Internet of Things (IoT)) [3], [4]. CR technology can be applied to the conventional wireless sensor networks, where there are many issues for research (e.g. channel handoff [5], [6], energy consumption [7], AMC [8], etc.). Especially, AMC is the core technique in CR networks to identify modulation types of unknown signals without prior knowledge. Normally, AMC

technique is mainly divided into two approaches, which are likelihood-based (LB) and feature-based (FB) approaches. Though the optimal classification performance is gained, LB approach has the significant drawbacks, which are the high computational complexity and the implement problem for the real-time systems [9]. On the contrary, FB approach can make the implement simple and gain suboptimal performance with low computational complexity [10]. The conventional FB approach is divided into a feature extraction part and a classification part, where in the feature extraction part instantaneous feature, High-Order Cumulant (HOC) feature [11], cyclostationary feature [12] and wavelet features [13] are mainly exploited [14], and in the classification part Support Vector Machine (SVM), Decision Tree (DT), k-Nearest Neighbor (kNN), etc. are generally adopted [15], [16]. On the contrary, the methods have limited classification

capability depending on the handcrafted feature as well as the HOC and cyclostationary methods may cause high computational complexity.

Recently, the deep learning-based techniques such as a CNN and a recurrent neural network (RNN) contain both the feature extraction part and the classification part, which has stronger representative capability than the classical FB approach [16], [17]. The deep learning techniques have been widely applied to diverse applications including AMC. Many AMC works based on the CNN focus on classification accuracy performance for analog and digital modulations (e.g. Amplitude Modulation (AM), M-ary Phase Shift Keying (PSK), M-ary Quadrature Amplitude Modulation (QAM) etc.), and show the competive performance [18]–[20]. However, one of the research works only considers a small number of modulation types, which is easier to achieve high classification performance. In another case, a simple channel condition for the wireless communication such as an additive white Gaussian noise (AWGN) channel is just considered, which can obtain relatively the clear feature compared with the Rayleigh fading channel. Yang *et al.* [21] represent the classification performance compared with the AWGN channel and the Rayliegh fading channel conditions. The result clearly shows that the performance is better in the AWGN channel. In terms of computational complexity, a image-based classification in the CNN such as a constellation diagram may cause high computational complexity due to three dimensions of the image [19]. Thus, many research works Lin *et al.* [22], Zhang *et al.* [23], [24], Hermavan *et al.* [25], represent the AMC performance by using the complex signal. In [22] HybridNet is proposed to show the enhanced performance, which exploits both CNN and a bidirectional gated recurrent unit to capture temporal depecdencies. In [23] a CNN-based AMC method is proposed, where the architecture is designed to improve the generalized capability under varying noise conditions. In [24] a multi-stream CNN is proposed, which shows the network architecture is extended horizontally to extract diverse key features and to mitigate the over-fitting problem. In [25] IC-AMCNet is proposed for beyond 5G communication (B5C), which is designed to consider both the predicted accuracy and the processing time of B5C requirement below 0.01 *ms*. However, all the related works keep the conventional input, which is $2 \times N$ IQ components. Thus, it may be limited to extract deep features from the conventional input.

In this paper, we propose a robust CNN architecture via a novel method. For the simulation the DeepSig: RadioML 2018.01A dataset including 24 modulation samples is utilized, which can be categorized into the difficult class and the normal class. The proposed method shows the enhanced classification performance compared with the conventional CNN models. The contributions of this paper are sumarized as follows:
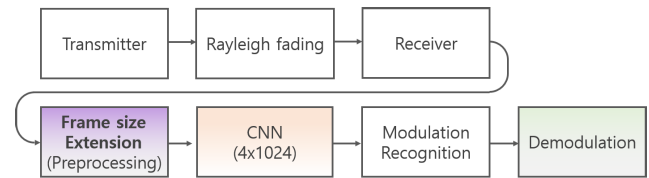

FIGURE 1. Proposed system model.

- A novel method extending the frame size from 2×1024 to 4×1024 by copying the own frame helps the networks extact deep features from the extended frame.
- A new model is proposed, which considers both the predicted accuracy and the computational complexity by reducing the extended input size with a key averagepooling layer.
- The accuracy performance of the proposed model outperforms the conventional models, which achieves 2.83% improvement at SNR 10 dB compared with the latest CNN model [26].

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. SYSTEM MODEL

The system model for the AMC process between the transmitter and the receiver is represented by Fig. 1, where assuming that the transmission signal is generated by a multi-input multi-output (MIMO) system, and the signal is impaired by the clockoffset from the local oscillator and by the Rayleigh fading channel. The received signal at the k-th observation can be expressed as

$$\mathbf{Y}_k = \mathbf{H}\mathbf{X}_k + \mathbf{N}. \tag{1}$$

where $\mathbf{Y}_k$ is $[y_1, y_2, \ldots, y_{m_r}]$, $\mathbf{H}$ is the Rayleigh channel matrix $m_r \times n_t$ where m_r is the receiver antenna number and n_t is the transmitter antenna number, $\mathbf{X}_k$ is the modulated symbol vector $[x_1, x_2, \ldots, x_{n_t}]^T$ and N is the AWGN. One of y_{m_r} can be expressed by

$$y_{m_r} = \beta e^{j(2\pi f_\Delta t + \phi_\Delta)} x_k + n. \tag{2}$$

where β is the multi-path amplitude, f_Δ is the carrier frequency and ϕ_Δ is the phase offset, respectively. The clockoffset by local oscillators and Doppler shift from Rayleigh fading channel cause the offsets. Above all, in the preprocessing step the received signals are aggregated to the frame length where the vector s is $[s_1, s_2, \ldots, s_N]$, $N = 1024$. Secondly, normalization is performed by root mean square (RMS) to help the networks quickly reach the global optimum point, which can be represented by

$$\bar{s}_i = \frac{s_i}{\sqrt{\frac{1}{N} \sum_{i=0}^{N-1} |s_i|^2}}. \tag{3}$$
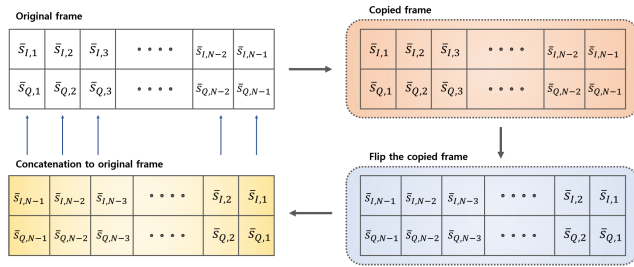
FIGURE 2. Process of frame extension.

The frame after normalization is divided into IQ components as the $Re[\bar{s}_i]$ term and the $Im[\bar{s}_i]$ term, which is equal to a $2 \times N$ matrix. In addition, To extend the $2 \times N$ matrix a copy version of the matrix is generated and before concatenate it below the original matrix, it will be horizontally flipped, which can be represented by

$$S_j = \begin{bmatrix} \bar{s}_{I,0} & \bar{s}_{I,1} & \cdots & \bar{s}_{I,N-1} \\ \bar{s}_{Q,0} & \bar{s}_{Q,1} & \cdots & \bar{s}_{Q,N-1} \\ \bar{s}_{I,N-1} & \bar{s}_{I,N-2} & \cdots & \bar{s}_{I,0} \\ \bar{s}_{Q,N-1} & \bar{s}_{Q,N-2} & \cdots & \bar{s}_{Q,0} \end{bmatrix}, \quad (4)$$

where I is the in-phase term and Q is the quadrature term. The high-dimension representation of the frame is equal to 4×1024, which is used as a input for the CNN model. Many of the samples in the frame have high-impact features, which are obvious features to represent each modulation type. Thus, This approach can obtain more high-impact features extracted through 3×1 convolutional operation and 2×1 average pooling operation. In addition, the expanded frame is simply implemented without high cost, but it can improve the accuracy. This concept may have trade-off between classification accuracy and computational complexity, which should be considered.

B. DATASET DESCRIPTION
In this paper, DeepSig:RadioML 2018.01A dataset [27] is utilized for the simulation, which is generated by synthesizing simulated signals with virltual propagation effects and captured signals with real propagation effects in industrial, scientific, and medical (ISM) band. The SNR range of the collected data is from $-10\ dB$ to $20\ dB$ with the step size of $2\ dB$, where 98304 frames are stored in each SNR according to the number of modulation types. Finally, the dataset is divided into 80% for training and the remaining for testing.

III. PROPOSED CNN MODEL
A. CNN MODEL
The CNN is one of advanced neural networks with hidden layers, which can recognize patterns by learning deeply. In general, to extract feature maps in CNN convolution, activation and pooling layers are mainly utilized. Firstly, in the convolution layer convolutional operation is performed by using convolutional filters to make arbitrary features, which
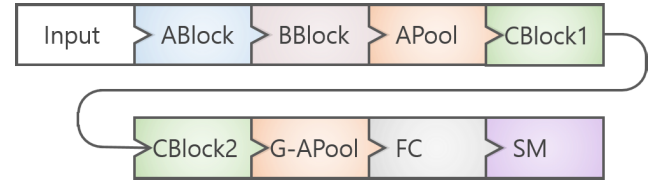
can be shown by

$$z_p = \sum_p V(p) \times W(p) + b, \quad (5)$$

where $V(p)$ is the input size with the two dimensions, $W(p)$ is the two-dimensional filter and b is the scalar bias, respectively. Secondly, the role of activation layer is to activate the features from the convolutional layer and the activated features are carried to the next layer. For activation functions there are several schemes such as sigmoid, hyperbolic tangent, rectified linear unit (ReLU), where ReLU is mainly used, which can be expressed by

$$\bar{z}_p = \begin{cases} z_p & \text{if } z_p \geq 0, \\ 0 & \text{if } z_p < 0, \end{cases} \quad (6)$$

In the pooling layer a down-sampling function can mitigate overfitting problem and decrease computational complexity by reducing the feature map volume, where max-pooling function as $max(\bar{z}_p)$ or average-pooling function as $mean(\bar{z}_p)$ according to the arbitrary filter size is generally used. In addition, as a loss function stochastic gradient descent with momentum (SGDM) optimizer is used in order to compute the gradients and update the CNN learnable parameters (i.e. weights and biases) based on the input.

B. PROPOSED CNN MODEL
The proposed CNN model is designed efficiently according to the input size for classification of 24 modulations, where the overall architecture of the proposed CNN is displayed in Fig. 3. To extract the feature maps a a-type block (ABlock), a b-type block (BBlock) and two c-type blocks (CBlocks) play a significant role, which contain multiple convolutional layers. Above all, in the ABlock the main function is downsampling of the input size performed by max-pooling layers (MPool) without loss of key features. The difference of the BBlock and the CBlock is that one convolutional layer with the 3×1 kernal is added in the BBlock, which can help the vertical features on the spatial dimension be well extracted. In addition, skip connections are deployed in the BBlock and the CBlock to mitigate the vanishing gradient problem, and all the block architectures are shown in Fig. 4. For the convolutional operation throughout the proposed CNN the convolution block (ConvB) is used, which contatins a convolutional layer, a batch normalization layer and a ReLU activation layer shown in Fig. 4c, where the stride is 1×1 in the convolutional layer. The point-wise convolution

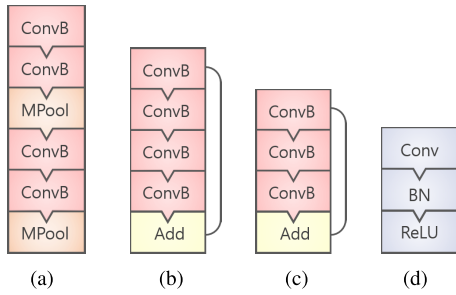**FIGURE 3.** Proposed CNN architecture.

FIGURE 4. Description of blocks in the proposed CNN. (a) ABlock; (b) BBlock; (c) CBlock1, 2; and (d) ConvB.

operation represents a linear combination of the output feature maps with 1×1 kernels, which can get the effect of dimensionality reduction. Accordingly, computational complexity decreases through the point-wise convolution. And the asymmetric kernels can minimize trainable parameters without accuracy performance degradation compared with 3×3 kernels. For example, in the BBlock the asymmetric kernels are used to extract features, which can reduce the trainable parameters by approximately half compared with 3×3 kernals. Regarding the impact of accuracy performance the literatures [27], [29] are referenced. In terms of computational complexity, a average-pooling layer (APool) between the BBlock and CBlock1 plays the key role to diminish the computational complexity due to the copied IQ samples as well as makes a impact for the classification accuracy. The reason is that the average-pooling operation can compress 4 rows to 2 rows while keeping the significant information from the input. In other points for reduction of computational complexity, point-wise convolutional layers are deployed in the BBlock and CBlock and asymmetric filters are equipped in the remain convolutional layers, which can reduce the substantial number of trainable parameters. The feature map size from the CBlock2 is completely reduced to 1×1 by a global average-pooling layer (G-APool) and the output is passed to a fully connected layer (FC). Finally, a softmax layer (SM) computes the probability for the classification. The detailed configuration of the proposed CNN architecture is summarized in Table 1.

IV. NUMERICAL RESULTS

The results of simulation works with the 24-modulation dataset are represented to demonstrate the efficiency of the novel method that is to extend the own frame from the receiver, and the robustness of the proposed CNN. For the simulation parameter configuration we set the max epochs to 45, the mini-batch size to 64, the initial learning rate to 0.1, the drop period of the learning rate to 20, the drop factor of the learning rate to 0.1 and the SGDM optimizer is applied to optimize the loss value, which is sumarized in Table 2. For training process 80% of the dataset is utilized and for testing the other is utilized, which is simulated by Matlab 2020b. The hardware specification for the simulation consists of an i5 2.9 GHz CPU, 32 GB RAM, and NVIDIA

TABLE 1. Configuration of the proposed CNN.

Type	Output Size	Filter Size	Parameters
Input	4×1024	-	-
ABlock	$4 \times 1024 \times 16$	$16 \times 1 \times 3$	96
	$4 \times 1024 \times 16$	$16 \times 3 \times 1$	806
	$4 \times 512 \times 16$	1×2	-
	$4 \times 512 \times 16$	$16 \times 1 \times 3$	806
	$4 \times 512 \times 16$	$16 \times 3 \times 1$	806
	$4 \times 256 \times 16$	1×2	-
BBlock	$4 \times 256 \times 32$	$32 \times 1 \times 1$	608
	$4 \times 256 \times 16$	$16 \times 1 \times 3$	1584
	$4 \times 256 \times 16$	$16 \times 3 \times 1$	816
	$4 \times 256 \times 32$	$32 \times 1 \times 1$	608
APool	$2 \times 256 \times 32$	2×1	-
CBlock1	$2 \times 256 \times 64$	$64 \times 1 \times 1$	2240
	$2 \times 256 \times 24$	$24 \times 1 \times 3$	4680
	$2 \times 256 \times 64$	$64 \times 1 \times 1$	1728
CBlock2	$2 \times 256 \times 128$	$128 \times 1 \times 1$	8576
	$2 \times 256 \times 32$	$32 \times 1 \times 3$	12384
	$2 \times 256 \times 128$	$128 \times 1 \times 1$	4480
G-APool	$1 \times 1 \times 128$	-	-
FC	$1 \times 1 \times 24$	-	3096
Softmax	24	-	-

TABLE 2. Configuration for the simulation.

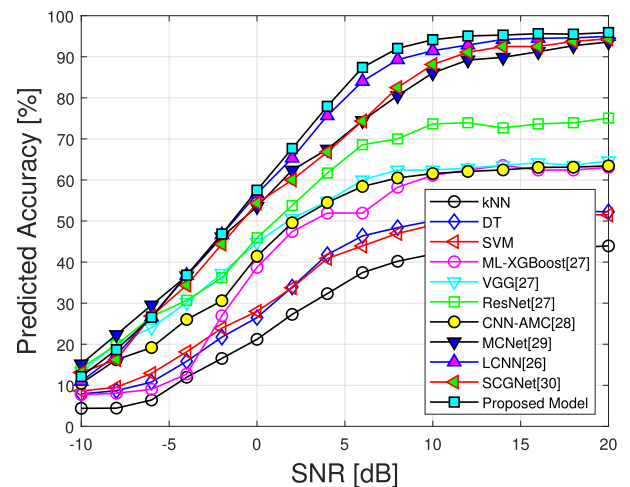| Type	Value
MaxEpochs	45
MiniBatchSize	64
InitialLearningRate	0.1
LearningRateDropPeriod	20
LearningRateDropFactor	0.1
Optimizer	SGDM

FIGURE 5. The accuracy evaluation with different models.

GeForce RTX 2080 Super GPU devices. To show the performance in terms of the classification accuracy and the
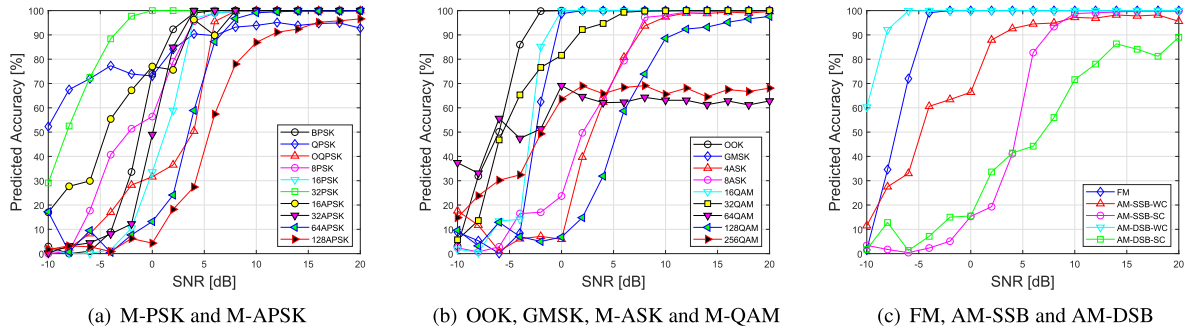
(a) M-PSK and M-APSK

(b) OOK, GMSK, M-ASK and M-QAM

(c) FM, AM-SSB and AM-DSB

FIGURE 6. Classification accuracy of 24 modulation types.

computational complexity the proposed model is compared with the related works that includes latest models such as MCNet [29], LCNN [26] and SCGNet [30].

First of all, the performance evaluation of the proposd model is represented with respect to the conventional models shown in Fig. 5. To accurately evaluate the performance the approaches are categorized into machine learning techniques such as kNN, DT, and SVM and deep learning technique. As shown, the accuracy performance of the machine learning techniques is lower in the most SNRs than deep learning technique. With respect to the deep learning approach, the performance of VGG [27] and CNN-AMC [28] is shown as a similar trend, which is the worst group. ResNet [27] equipped with the skip connections largely outperforms the worst group in the high SNRs (i.e. over 10 dB) even though the fundamental structure is almost same as VGG. Regarding the best group including the latest models (i.e. MCNet, LCNN and SCGNet) the performance is more outstanding in the most SNRs than the aforementioned models. Above all, LCNN that is the top in the best group from −4 dB to 20 dB SNRs is designed for the light-weight structure, which presents 56.64% and 91.48% accuracy at 0 dB and 10 dB SNR, respectively. Finally, the proposed model outperforms the LCNN model by 1.51% and 2.83% at 0 dB and 10 dB SNR, respectively. Through the literatures [27], [29], [30] it is shown that it is hard to obtain better accuracy when it is converged near maximum in high SNRs even though it uses more trainable parameters. Thus, 2.83% improvement represents high improvement. According to the result the proposed model with the effective structure can extract deep features to achieve superior accuracy, which also proves that the extension frame can help the deep features be produced in the networks.

For the 24-modulation accuracy performance of the proposed model the results are shown in Fig. 6 in detail, which is divided into 3 groups. In the first group in Fig. 6a, from 8 dB to the end SNRs all the modulations reach the accuracy over 90% except 128 APSK even though the most modulations are difficult to predict especially in the bad channel conditions. On the contrary, from 6 dB to −6 dB the classification accuracy of the schemes at large decreases rapidly, where

64 APSK suffers from a significant performance degradation next to 128 APSK. Thus, 64 and 128 APSK are more vulnerable to predict as the SNRs weakens. In the second group in Fig. 6b, 64 and 256 QAM keep the accuracy relatively low as the SNR increases upward due to the lack of identical features, which is major concern to solve. On the other hand, OOK and GMSK that are the low order modulations easily reach high accuracy in low SNR, which yield 100% and 98.7% at 0 dB SNR, respectively. In the last group in Fig. 6c the difficult modulation type is AM-DSB-SC, which rarely reach over 90% in the high SNRs even though the remains easliy keep over 90%. However, as the SNR decreases to the low domain the performance of AM-SSB-SC drops rapidly. Finally, both AM-SSB-SC and AM-DSB-SC present similar performance trends in the low SNRs. In low SNRs region there are large fluctuations (e.g. the QPSK curve in Fig. 6(a), the 64 QAM curve in Fig. 6(b) and the AM-DSB-SC curve in Fig. 6(c)). Those are caused by the lightweight and simple CNN architecture, which can be relatively weak to extract the deep features for identification of each modulation in low SNRs. Moreover, QPSK, 64 QAM and AM-DSB-SC signals have the similar patterns with 256 QAM, AM-SSB-WC and 128 QAM, respectively. Thus, when the high-level features are not extracted, the fluctuation phenomenon can occur.

Based on the aforementioned results, for a visual analysis a confusion matrix is presented at 10 dB SNR in Fig. 7. According to the confusion matrix, most of the modulations are excellently recognized. However, it is known that AM-DSB-SC, 64 QAM and 256 QAM are difficult to recognize. The reason is that AM-DSB-SC is mainly confused with 128 APSK and 128 QAM, and 64 QAM and 256 QAM are mostly confused by AM-SSC-WC and QPSK, respectively.

The performance comparison to the frame sizes is shown in Table 3. The effect of the extension method is not significantly outstanding when compared with the no extension frame. However, in the high SNR region the extension frame can help the networks extract more features. According to the result the 4 × 1024 frame shows that the performance is approximately 1% on average from 6 dB to 20 dB SNR better than the 2 × 1024 frame. The reason why the extended frame method shows weak robustness in the low SNRs is
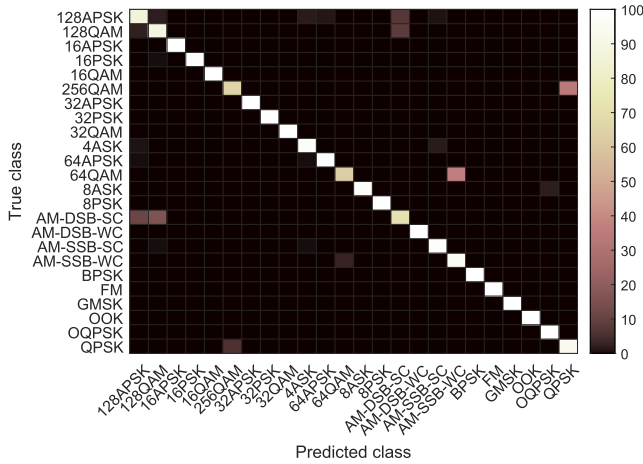
FIGURE 7. Confusion matrix of the proposed model at SNR 10 dB, where the classification accuracy is 94.15%.

FIGURE 8. Computational complexity analysis.

TABLE 3. Performance comparison to frame sizes.

SNR	2 × 1024	4 × 1024	6 × 1024	2 × 2048
-10	9.13	12.13	11.96	9.71
-8	18.02	18.66	18.71	15.77
-6	25.89	26.56	26.86	25.85
-4	35.83	36.82	35.91	36.09
-2	46.46	46.85	46.59	47.43
0	57.36	57.51	56.69	57.34
2	67.08	67.67	66.99	67.76
4	77.48	77.93	77.46	77.78
6	86.34	87.4	86.26	86.15
8	90.44	92.04	91.15	91.65
10	92.41	94.15	93.95	93.14
12	93.95	95.07	94.8	94.58
14	94.46	95.28	93.95	94.58
16	94.7	95.62	95.08	94.75
18	94.75	95.5	95.08	94.94
20	95.15	95.9	95.76	95.59

TABLE 4. Performance comparison on RML2016.10B.

SNR	2 × 1024	4 × 1024
-10	42.13	43.54
-8	53.12	54.3
-6	65.91	66.66
-4	78.02	78.85
-2	86.12	86.93
0	90.7	91.35
2	92.28	92.86
4	92.93	93.49
6	93.12	93.75
8	92.09	93.1
10	92.51	93.27
12	92.44	92.96
14	92.62	93.13
16	92.18	93.01
18	92.47	93.08

the frames are impaired by the effects of Rayleigh fading and thermal noise. Thus, it is relatively difficult to overcome the impairment due to low signal strength and extract the discriminative features for each modulation even though the frames are extended. On the contrary, in the high SNRs the extended frames can help the discriminative features extracted by the proposed model. For example, at the first index of the extended frame two in-phase, quadrature signal values are positioned, which are similar but different. Thus, high-impact features to represent the modulation identification can be strengthened by 3×1 convolution and 2×1 pooling operations. In addition, the 6×1024 frame is combined with 4×1024 frame and 2×1024 frame, where 2×1024 frame is aligned in random order. The 2×2048 frame is composed by horizontally concatenating the 2×1024 frame with the 2×1024 extension frame.

In Table 4. the additional performance comparison with 2*128 and 4*128 which is the extension version is shown to represent the effect of the proposed method through a new
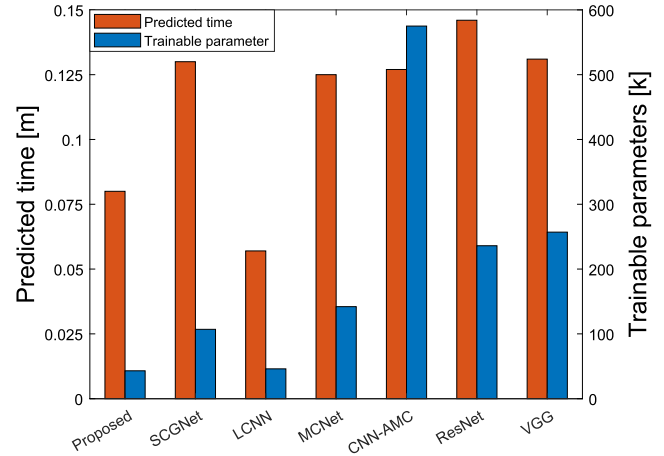
dataset, RML2016.10b [31]. The dataset has 10 modulation types, which are 8 PSK, AM-DSB, BPSK, CPFSK, GFSK, PAM4, 16 QAM, 64 QAM, QPSK and WBFM, respectively. Each type contains the number of 90,000 frames from -10 dB to 18 dB. Therefore, 900,000 frames are used for the simulation at the same condition as the RML2018.10A dataset. As shown in Table 4, the performance of the proposed method is better in overall SNRs, which shows almost the same result as the RadioML 2018.01A dataset. Therefore, the proposed method represents the valid effect based on the simulation results.

To evaluate the computational complexity, Fig. 8 presents it regarding the seven CNN models. As a result, the predicted time of the proposed model is slightly longer than the LCNN by approximately 0.03 *ms* even though the trainable parameter of the proposed model is smaller by roughly 6.5%. The reason is that the input size increases by the extension frame. The extended frame may cause more computational complexity compared with no extended frame method as displayed in Fig. 8. The predicted time of the proposed

model, SCGNet, LCNN, MCNet, CNN-AMC, ResNet and VGG takes 0.08, 0.13, 0.057, 0.125, 0.127, 0.146 and 0.131 millisecond, respectively. Considering that the proposed model is similar with the LCNN model, the proposed method can increase more by 0.023 millisecond additionally. However, comprehensively the impact compared with the others is not fatal. Thus, it can be applied to a general system model. The predicted time of the other models is quite longer because SCGNet, MCNet and ResNet adopt multiple blocks which causes additional operations and deep convolutional networks. Further, VGG structure is not efficient for the lightweight due to 3 × 3 kernels and CNN-AMC adopts several FC layers, which can cause considerable trainable parameters. Comprehensively, the proposed model applying the novel method presents the robust performance compared with the others despite the slight increase of the predicted time. Thus, the novel method causes a tradeoff between the accuracy and computational complexity.

V. CONCLUSION

In this paper, we propose a new method that extends the own frame by copying and flipping itself, which can improve the recognition accuracy even though the extended frame is not original from the transmitter. According to the input size the CNN architecture for AMC is effectively designed, where point-wise convolutions and asymmetric kernels are adopted to reduce the computational complexity. For the simulation evaluation the DeepSig:RadioML 2018.01A dataset is utilized, which contains 24 modulation types. According to the simulation results the classification accuracy of the proposed model is superior to the other models in the SNR range from -4 *dB* to 20 *dB*, where the accuracy performance is 94.15% at 10 *dB* SNR. In terms of the computational complexity, the predicted time of the proposed model is slightly longer than the LCNN by approximately 0.03*ms* due to the increase of the input size. However, comparing with the other models the proposed model gets relatively low complexity. Therefore, the robustness of the proposed model is verified via a comparison with the state-of-the-art models in terms of the accuracy and complexity performance. One of the shortcomings of the proposed method is that in low SNR region it can not help the discriminative features extracted due to the low signal strength and the impairment. The other is that it causes more computational complexity compared with the normal frame method. The proposed method has very low cost to apply to a system and low complexity to implement. Thus, for the future works the comparison with the simulation and the implement performance is performed to evaluate the effects of the proposed method in a real system.

REFERENCES

[1] M. Luis, R. Oliveira, R. Dinis, and L. Bernardo, "RF-spectrum opportunities for cognitive radio networks operating over GSM channels," *IEEE Trans. Cognit. Commun. Netw.*, vol. 3, no. 4, pp. 731–739, Dec. 2017.

[2] 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.

[3] J. Lunden, S. R. Kulkarni, V. Koivunen, and H. V. Poor, "Multiagent reinforcement learning based spectrum sensing policies for cognitive radio networks," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 5, pp. 858–868, Oct. 2013.

[4] 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.

[5] S. Haykin and P. Setoodeh, "Cognitive radio networks: The spectrum supply chain paradigm," *IEEE Trans. Cognit. Commun. Netw.*, vol. 1, no. 1, pp. 3–28, Mar. 2015.

[6] A. O. Bicen, E. B. Pehlivanoglu, S. Galmes, and O. B. Akan, "Dedicated radio utilization for spectrum handoff and efficiency in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 9, pp. 5251–5259, Sep. 2015.

[7] G. I. Tsiropoulos, O. A. Dobre, M. H. Ahmed, and K. E. Baddour, "Radio resource allocation techniques for efficient spectrum access in cognitive radio networks," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 824–847, 1st Quart., 2016.

[8] 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.

[9] 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.

[10] S. Majhi, R. Gupta, W. Xiang, and S. Glisic, "Hierarchical hypothesis and feature-based blind modulation classification for linearly modulated signals," *IEEE Trans. Veh. Technol.*, vol. 66, no. 12, pp. 11057–11069, Dec. 2017.

[11] A. Swami and B. M. Sadler, "Hierarchical digital modulation classification using cumulants," *IEEE Trans. Commun.*, vol. 48, no. 3, Mar. 2000, Art. no. 416429.

[12] 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.

[13] G. Latif, D. N. F. A. Iskandar, J. M. Alghazo, and N. Mohammad, "Enhanced MR image classification using hybrid statistical and wavelets features," *IEEE Access*, vol. 7, pp. 9634–9644, Dec. 2019.

[14] 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, Aug. 2017.

[15] K. Zhang, E. L. Xu, H. Zhang, Z. Feng, and S. Cui, "Data driven automatic modulation classification via dictionary learning," *IEEE Wireless Commun. Lett.*, vol. 7, no. 4, pp. 586–589, Aug. 2018.

[16] M. Elwekeil, S. Jiang, T. Wang, and S. Zhang, "Deep convolutional neural networks for link adaptations in MIMO-OFDM wireless systems," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 665–668, Jun. 2019.

[17] 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.

[18] S. Huang, Y. Jiang, Y. Gao, Z. Feng, and P. Zhang, "Automatic modulation classification using contrastive fully convolutional network," *IEEE Wireless Commun. Lett.*, vol. 8, no. 4, pp. 1044–1047, Aug. 2019.

[19] 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.

[20] S. Zheng, P. Qi, S. Chen, and X. Yang, "Fusion methods for CNN-based automatic modulation classification," *IEEE Access*, vol. 7, pp. 66496–66504, May 2019.

[21] 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, Aug. 2019.

[22] R. Lin, W. Ren, X. Sun, Z. Yang, and K. Fu, "A hybrid neural network for fast automatic modulation classification," *IEEE Access*, vol. 8, pp. 130314–130322, Jul. 2020.

[23] T. Zhang, C. Shuai, and Y. Zhou, "Deep learning for robust automatic modulation recognition method for IoT applications," *IEEE Access*, vol. 8, pp. 117689–117697, Mar. 2020.

[24] 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, Feb. 2020.

[25] 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.

[26] S.-H. Kim, J.-W. Kim, V.-S. Doan, and D.-S. Kim, "Lightweight deep learning model for automatic modulation classification in cognitive radio networks," *IEEE Access*, vol. 8, pp. 197532–197541, Oct. 2020.

[27] 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.

[28] 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.

[29] 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.

[30] G. B. Tunze, T. Huynh-The, J.-M. Lee, and D.-S. Kim, "Sparsely connected CNN for efficient automatic modulation recognition," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 15557–15568, Dec. 2020.

[31] Deepsig. (2016). *RF Datasets for Machine Learning*. [Online]. Available: https://www.deepsig.ai/datasets

WILLIAMS-PAUL NWADIUGWU (Member, IEEE) received the Ph.D. degree in IT convergence engineering from the Kumoh National Institute of Technology (KIT), Republic of Korea, in 2020. He is currently working as a Senior Research Scientist and an Associate Research Professor with the ICT-Convergence Research Center, School of Electronic Engineering, Kumoh National Institute of Technology (KIT). He has authored over 25 scholarly articles/conference papers and SCI/E journals combined till date. His major and current research interests include solving real-time reliability constraints in the industrial IoT systems, wireless AG2UG/UG2AG/UG2UG communications, robotics and mobility sensor networks, missile guidance and control, and ultrawideband (UWB) network embedded systems. He is a Registered Member of the IEEE Board.
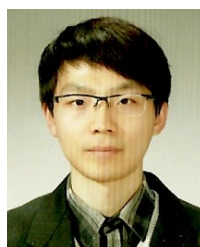
SEUNG-HWAN KIM (Member, IEEE) received the B.E., M.S., and Ph.D. degrees in electronics engineering from the Kumoh National Institute of Technology (KIT), in 2010, 2012, and 2018, respectively. He is currently working with the ICT Convergence Research Center, Kumoh National Institute of Technology, Republic of Korea, as a Senior Researcher. His research interests include 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 with Optical Communications Company Artech, from 2014 to 2018. He is currently working with the ICT Convergence Research Center, Kumoh National Institute of Technology, Republic of 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.

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

● ● ●